Robinson's Arithmetic

We're developing the idea that a set S is Σ iff it's effectively enumerable iff there is a proof procedure for S. We now want to see that we can we can take the notion of "proof procedure" literally, by treating a proof procedure as a derivation within a certain system of axioms. So we now need to look at systems of axioms.

Definition. *Q*, also known as *Robinson's arithmetic*, is the conjunction of the following axioms:

(Q1) $(\forall x)\neg\, x = 0$

(Q2) $(\forall x)(\forall y)(sx = sy \rightarrow x = y)$

(Q3) $(\forall x)((x + 0) = x$

(Q4) $(\forall x)(\forall y)(x + sy) = s(x + y)$

(Q5) $(\forall x)(x \bullet 0) = 0$

(Q6) $(\forall x)(\forall y)(x \bullet sy) = ((x \bullet y) + x)$

(Q7) $(\forall x)(xE0) = s0$

(Q8) $(\forall x)(\forall y)(xEsy) = ((xEy) \bullet x)$

(Q9) $(\forall x)\neg\, x < 0$

(Q10) $(\forall x)(\forall y)(x < sy \leftrightarrow (x < y \vee x = y))$

(Q11) $(\forall x)(\forall y)(x < y \vee (x = y \vee y < x))$

As an account of the natural numbers, Q is pitifully weak. Even the very simplest generalizations, like the commutation law of addition and the commutative law of multiplication, are underivable in Q. Nevertheless, we have the following:

Theorem. Every true Σ sentence is derivable in Q.

This theorem is why Q is worth looking at. Q is of no interest in itself. Our reason for bringing it up is that it's a single-axiom theory within which every true Σ sentence is provable.

Proof: First, note that, for each m and n,

$$0 = [0]$$

$$s[m] = [sm]$$

$$([m] + [n]) = [m+n]$$

$$([m] \bullet [n]) = [m \bullet n]$$

$$([m]E[n]) = [mEn]$$

are all consequences of Q. An easy induction on the complexity of terms then enables us to prove that, for each closed term τ, there is a number n such the sentence

$$\tau = [n]$$

is a consequence of Q. An induction shows that each number m has this property:[1]

$$(\forall n)(m \neq n \rightarrow Q \vdash \neg [m] = [n]$$

A similar induction shows that, for each number n, we have:

For every m, if m < n, then [m] < [n] is provable in Q, whereas, if m ≥ n, then [m] < [n] is refutable[2] in Q.

Thus we see that every atomic sentence is decidable[2] in Q. It follows immediately that every quantifier-free sentence is decidable in Q. Because

$$Q \vdash (\forall x)\neg x < 0$$

---

1 "Γ ├ ϕ" means that ϕ is a consequence of Γ.

2 A sentence is *refutable* in Q iff its negation is provable in Q. A sentence is *decidable* in Q iff it is either provable or refutable.

and, for each n,

$$Q \vdash (\forall x)(x < [n+1] \leftrightarrow (x = [0] \lor x = [1] \lor \ldots \lor x = [n])),$$

every bounded formula is provably equivalent to an quantifier-free formula. We eliminate bounded quantifiers from he outside in, just as before.

We now see that every bounded sentence is decidable in Q, and so, since Q is true, every true bounded sentence is provable in Q. Consequently, every true Σ sentence can be proven by providing a witness.⊠

Corollary. Let Γ be a true theory that includes[3] Q. Then for each Σ set[4] S, there is a Σ formula that weakly represents S in Γ.

Proof: Let S be the extension of the Σ formula ϕ. If n is in S, ϕ([n]) is a consequence of Q, and so a consequence of Γ. If n ∉ S, ϕ([n]) isn't true, and so it isn't a consequence of Γ.⊠

We can strengthen this corollary by employing a new notion:

Definition. A theory Γ is ω-*inconsistent* iff, for some formula ψ(x), Γ proves (∃x)ψ(x), but it also proves ¬ψ([n]), for each n.

Since an inconsistent theory proves every sentence, every inconsistent theory is ω-inconsistent, but, as we shall see later, not every ω-inconsistent theory is inconsistent. Every true theory is ω-consistent, but not every ω-consistent theory is true.

---

3 To say that Γ *includes* Q, in standard usage, it's not literally required that Q be an element of Γ. It's enough that Q is a consequence of Γ. The trouble is that, in standard usage, "theory" is ambiguous between a set of axioms and the set of consequences of the set of axioms. The ambiguous usage is thoroughly entrenched, so we have to live with it.

4 As usual, what we say about sets goes for relations too.

Corollary. Let Γ be an ω-consistent theory that includes[5] Q. Then for each Σ set S, there is a Σ formula that weakly represents S in Γ.

Proof: Let S be the extension of (∃y)ψ(x,y), where ψ is bounded. The argument that, if n is in S, then Γ ├ (∃y)ψ([n],y), is the same as above. If n isn't in S, then, for each m, ψ([n],[m]) is false, and so ¬ψ([n],[m]) is a consequence of Q, and hence a consequence of Γ. It follows by ω-consistency that (∃y)ψ([n],y) isn't a consequence of Γ.⊠

We cannot strengthen the corollary still further by replacing "ω-consistent" by "consistent," for it is possible to find a consistent theory that includes Q in which not every Σ set is weakly representable. The proof proceeds by starting with a set K that is Σ but not Δ, and by enumerating all the formulas with one free variable. We build up our theory Γ in stages, starting with Q, and at the nth stage adding a sentence to the theory that kills off the possibility that the nth formula weakly represents K, maintaining consistency all the while. I won't go into details.

One can, however, show that, if Γ is a consistent, Σ set of sentences that implies Q, then every Σ set is weakly representable in Γ. The proof requires machinery we haven't developed yet.

Theorem (Rosser). For any Δ set S, there is a Σ formula that strongly represents S in any consistent theory that includes Q.

---

5 To say that Γ *includes* Q, in standard usage, it's not literally required that Q be an element of Γ. It's enough that Q is a consequence of Γ. The trouble is that, in standard usage, "theory" is ambiguous between a set of axioms and the set of consequences of the set of axioms. The ambiguous usage is thoroughly entrenched, so we have to live with it.

Proof: If S is Δ, then there are bounded formulas ϕ(x,y) and ψ(x,y) such that (∃y)ϕ(x,y) weakly represents S in Q and (∃y)ψ(x,y) weakly represents the complement of S. We want to put these formulas together to construct a single formula such that the formula weakly represents S in Q and its negation weakly represents the complement of S. If we were working with true arithmetic rather than Q, we could just take our formula to be (∃y)ϕ(x,y), taking advantage of the fact that (∀x)(¬(∃y)ϕ(x,y) ↔ (∃y)ψ(x,y)) is true. However, we are working with Q, and (∀x)(¬(∃y)ϕ(x,y) ↔ (∃y)ψ(x,y)), though true, might not be provable in Q. So we have to be more devious.

The way our formula θ(x) is constructed is reminiscent of the way we proved the Reduction Theorem for effectively enumerable sets. There we had effectively enumerable sets A and B, and we wanted to find nonoverlapping effectively enumerable sets C ⊆ A and D ⊆ B with C ∪ D = A ∪ B. The idea was to simultaneously list A and B. If n first turns up in the list for A, put n into A, whereas if n first turns up in the list for B, put it in D; ties go to C. The formula θ(x) that we're trying to produce describes an analogous construction in which, given n, we simultaneously try to construct a witness to (∃y)ϕ([n],y) and to construct a witness to (∃y)ψ([n],y). If our first witness is a witness to (∃y)ϕ([n],y), make θ([n]) true, whereas if our first witness is a witness to (∃y)ψ([n],y), make θ([n]) false; ties go to truth.

The little parable I just told isn't part of the proof. The proof consists in writing down a formula and verifying that its works. The parable was intended to motivate the choice of formula. Whether of not the parable worked, here is our formula θ(x):

$$(\exists y)(\phi(x,y) \wedge (\forall z < y)\neg\, \psi(x,y)).$$

Let Γ be a consistent theory that includes Q. We need to verify the following four statements:

(a) If n is in S, then $\Gamma \vdash \theta([n])$.

(b) If n isn't in S, then $\Gamma \vdash \neg\theta([n])$.

(c) If n is in S, then $\Gamma \nvdash \neg\theta([n])$.

(d) If n isn't in S, then $\Gamma \nvdash \theta([n])$.

Proof of (a): In n is in S, then $\theta([n])$ is a true Σ sentence, provable in Q and hence in Γ.

Proof of (b): If n isn't in S, then, for some natural number m, $\psi([n],[m])$ is a true bounded sentence, and so a theorem of Q. Consequently,

(1) $(\forall y)([m] < y \rightarrow (\exists z < y)\psi([n],z))$

is a consequence of Q. So are

(2) $(\forall y)([m] < y \rightarrow \neg(\forall z < y)\neg \psi([n],z))$

and

(3) $(\forall y)([m] < y \rightarrow \neg(\phi([n],y) \wedge (\forall z < y)\neg \psi([n],z)))$.

Because n isn't in S, for each k, $\phi([n],[k])$ is false. Consequently, for each k, $\neg(\phi([n],[k]) \wedge (\forall z < [k])\neg \psi([n],z))$ is true. Therefore,

(4) $(\forall y)(y < [m] \rightarrow \neg(\phi([n],y) \wedge (\forall z < y)\neg \psi([n],z)))$

is a true bounded sentence, and so a consequence of Q. Also,

(5) $\neg(\phi([n],[m]) \wedge (\forall z < y)\neg \psi([n],z))$

is a true bounded sentence, and so a consequence of Q. (5) is equivalent to

(6) $(\forall y)([m] = y \rightarrow \neg(\phi([n],y) \wedge (\forall z < y)\neg \psi([n],z)))$.

(Q11) gives us this:

(7) $(\forall y)([m] < y \vee ([m] = y \vee y < [m]))$

Combining (3), (4), (6), and (7), we see that

(8) $(\forall y)\neg(\phi([n],y) \wedge (\forall z < y)\neg \psi([n],z))$,

which is equivalent to

(9) $\neg\theta([n])$,

is a the consequence of Q, and hence a consequence of Γ.

Proof of (c): If n is in S, then, by (a), $\Gamma \vdash \theta([n])$. It follows by consistency that $\Gamma \nvdash \neg\theta([n])$.

Proof of (d): If n isn't in S, then by (b), $\Gamma \vdash \neg \theta([n])$. It follows by consistency that $\Gamma \nvdash \theta([n])$.⊠

Definition. A formula σ(x,y) *functionally represents* a total function f in a theory Γ iff, for each n, the sentence $(\forall y)\sigma([n],y) \leftrightarrow y = [f(n)])$ is a consequence of Γ.

Notice that, if our theory Γ (which includes Q) is consistent, any formula that functionally represents f in Γ also strongly represents f in Γ. The converse doesn't hold, in general. If θ strongly represents f in Γ, then, for each m and n,

$(\theta([n],[m]) \leftrightarrow [m] = [f(n)]$

is a consequence of Γ. So we can prove each instance of the generalization:

$(\forall y)(\theta([n],y) \leftrightarrow y = [f(n\_]),$

but there isn't any way to put the proofs of the infinitely many instances together to get a proof of the generalization. So, whereas Rosser's result gives us, for each Δ total function f, a formula that strongly represents f, that formula does not, as a rule, also functionally represent f. However, we can find another formula that does functionally represents f, as we shall now see:

Theorem (Tarski, Mostowski, and Robinson). For any Σ total function f, there is a Σ formula that functionally represents S in any theory that includes Q.

Proof: Since any Σ total function is Δ, Rosser's result tells us that there is a Σ formula $\theta(x,y)$ that strongly represents f in Q. Let $\sigma(x,y)$ be the following formula:

$$(\theta(x,y) \wedge (\forall z < y)\neg\, \theta(x,z)).$$

The proof that σ functionally represents f in Q (and hence in any theory that includes Q) is a lot like the last proof. Take any n.

If $k < f(n)$, $Q \vdash \neg\, \theta([n],[k])$, and hence $Q \vdash \neg\, \sigma([n],[k])$. Also, $Q \vdash \neg\, [k] = [f(n)]$, and so $Q \vdash (\sigma([n],[k]) \leftrightarrow [k] = [f(n)])$. Since $(\forall y)(y < [f(n)] \rightarrow (\sigma([n],y) \leftrightarrow y = [f(n)]))$ is provably (in Q) equivalent to the conjunction of all the sentences of the form $(\sigma([n],[k]) \leftrightarrow [k] = [f(n)])$ with $k < f(n)$, we see that

(10) $$(\forall y)(y < [f(n)] \rightarrow (\sigma([n],y) \leftrightarrow y = [f(n)]))$$

is a theorem of Q.

Since $(\forall z < [f(m)])\, \neg\, \theta([n],z)$ is provably (in Q) equivalent to the conjunction of all the sentences of the form $\neg\, \theta([n],[k])$, with $k < f(n)$, and since, for each $k < f(n)$, $\neg\, \theta([n],[k])$ is a consequence of Q, $(\forall z < [f(m)])\neg\, \theta([n],z)$ is a consequence of Q. $\theta([n],[f(n])$ is likewise a consequence of Q, so that Q implies $\sigma([n],[f(n)])$, which is logically equivalent to this:

(11) $$(\forall y)(y = [f(n)] \rightarrow (\sigma([n],y) \leftrightarrow y = [f(n)])).$$

Since Q implies $\theta([n],[f(n)]$, it also implies

(12) $$(\forall y)([f(n)] < y \rightarrow (\exists z < y)\theta([n],z)).$$

(12) is logically equivalent to this:

(13) $$(\forall y)([f(n) < y \rightarrow \neg(\forall y)\neg\, \theta([n],z)),$$

which immediately implies this:

(14) $$(\forall y)([f(n) < y \rightarrow \neg(\theta([n],y) \wedge (\forall y)\neg\, \theta([n],y))),$$

that is,

(15) $(\forall y)([f(n)] < y \rightarrow \neg\, \sigma([n],y))$.

Also, because Q implies

(16) $\neg\, [f(n)] < [f(n)]$,

Q implies this:

(17) $(\forall y)([f(n)] < y \rightarrow \neg\, y = [f(n)])$.

(15) and (17) together imply this:

(18) $(\forall y)([f(n)] < y \rightarrow (\sigma([n],y) \leftrightarrow y = [f(n)]))$.

(10), (11), (18), and (Q11) together imply:

(19) $(\forall y)(\sigma([n],y) \leftrightarrow y = [f(n)])$.⊠

Robinson’s Arithmetic has no intrinsic interrest for us. It’s technically useful as a means of proving some theorems, but it’s not independently important. In particular, proofs in Q scarcely resemble our intuitive ways of thinking about the natural numbers. We now turn our attention to a much stronger theory, Peano Arithmetic, that does a very good job of reflecting the ways we reason when we prove things informally about the natural numbers.