1 Introduction

The existence of nonstandard models of arithmetic was discovered in the 1930's. Though Gödel somehow missed it, their existence follows from the Gödel Completeness Theorem (1930) or the later Compactness Theorem, and more interesting examples follow from Gödel's Incompleteness Theorem of 1931. Skolem (who had serious credentials in number theory) gave an explicit construction in 1934.

Our basic object of study is $\mathbb{N}$ $(= \{0, 1, 2, ...\})$, with the usual $+, \cdot, <$ as primitive notions. I prefer to work in the ring $\mathbb{Z}$. This is a commutative ring with 1, discretely ordered by $<$. [Later we spell out the formal axioms for DOR, the class of discretely ordered rings with 1. Usually R is such a ring, and $N = \{x \in R : x \geq 0\}$.]

<u>Full Induction</u>

This is of course the axiom: $X \subseteq N,\ 0 \in X,\ X$ closed under $t \to t+1 \Rightarrow X = N$.

We all know that $\mathbb{Z}$ is (up to isomorphism) the only DOR satisfying Full Induction. So our nonstandard models never satisfy Full Induction.

If we restrict the X in induction we can get nonstandard models. The most common restriction is to restrict X to be first-order definable from $+, \cdot, <$. Rather than going straight to the jungle of syntax (eventually we must) we define **first-order** thus (for $R \in DOR$) :

i) If $f \in R[x_1, ...x_n]$, $\{\bar{x} \in R^n : f(\bar{x}) = 0\}$ and $\{\bar{x} \in R^n : f(\bar{x}) > 0\}$ are first-order;

ii) Close under $\cap, \cup, \neg$, projection $R^{m+n} \to R^n$ (repeatedly) to get R-definable (or first-order definable) subset of R^n.

[Since $\forall \bar{y} = \neg\exists\bar{y}\neg$, one has closure under $\forall$.]

<u>Peano arithmetic (PA)</u> is (the axiomatization of) the class of DOR with induction for X first-order definable.

THERE ARE NONSTANDARD MODELS of PA (in abundance), and we will for now work in an arbitrary one. Later we discuss particularly striking ones, where one may have $f \in \mathbb{Z}[x_1, ...x_9]$ unsolvable in $\mathbb{Z}$, but solvable in R.

The conventional wisdom is that PA is **weak**, because (Gödel) incomplete. In fact it is extremely strong (for example, Fermat's Last Theorem holds in all models of PA), and one has to work very hard to find perspicuous number-theoretic statements independent of (i.e. unsettled by) PA (the best examples, due to Paris et al in the 1970's, involve subtle variants of Ramsey's Theorem in finite combinatorics).

Philosophical issue (Kreisel): If I know that Φ is true in all models of PA (and not merely in $\mathbb{Z}$), what more do I know about $\mathbb{Z}$ than if I merely know that Φ is true in $\mathbb{Z}$?

In fact there is much more information, and this information grows as we weaken induction (thereby increasing the class of DOR considered). This will be made more precise later.

2 Algebraic preliminaries

Some notions make immediate sense in R, and we can prove for them the basic things one knows for $\mathbb{Z}$.

Division: $x \mid y \Leftrightarrow (\exists t)(y = xt)$.
Unit: $x \mid 1$ (± 1 are only units.)
Irreducible: x irreducible $\Leftrightarrow (x = yz \rightarrow y$ or z a unit$)$

Note that all relations above are first-order. So one easily proves by induction:

x not a unit $\Rightarrow x$ is divisible by an irreducible > 1.

Note: The correct property to induct on is:

$\Phi(x) : (\forall t)(0 < t \leq x \rightarrow t = 1$ or t is divisible by an irreducible $> 1)$.

What one does **not** get is factorization into a **finite** product of irreducibles (more on this later).

Primes x is prime $\Leftrightarrow (x \mid yz \rightarrow x \mid y$ or $x \mid z)$. This is of course first-order definable. In all rings, prime $\Rightarrow$ irreducible.

In models of PA, irreducible $\Rightarrow$ prime, and we give the sketch of the proof. The crucial point is that for $a, b \in R$ there are $\lambda, \mu \in R$ so that $\lambda a + \mu b$ divides both a and b. For them, up to $\pm 1, \lambda a + \mu b$ is the greatest common divisor of a, b.

First we establish Euclidean division in N. Say $m, n > 0$. Then $\exists k, \ell \ \ m = k.n + \ell, \ 0 \leq \ell < n$. If $m < n$ clear. Otherwise, suppose $m \geq n$. Now $mn \geq n$, so by induction there must be k with $n.k \leq m < n.(k+1)$, so let $\ell = m - kn$. [Later we shall look at how much induction needed here.]

Now return to the $\lambda a + \mu b$ argument. We can assume $a, b \in N$, and we do induction on (a, b). (If $a = b$, result clear.) Then if $a < b$, write $b = ka + \ell$, $0 \leq \ell < a$, and work with (a, ℓ). Then get $\lambda a + \mu \ell$ dividing a, ℓ, so $\lambda a + (b - ka)\ell$ divides a, ℓ (whence a and $ka + \ell = b$), so $(\lambda - k\ell)a + b\ell$ works.

In purely algebraic terms this shows: R a model of PA $\Rightarrow$ all finitely generated ideals are principal. (such rings are called Bezout), and algebra gives: Bezout $\Rightarrow$ (irreducible = prime). [Again, we later analyze how much induction really needed.] Also, p prime $\Rightarrow$ $R/(p)$ a field. The usual primes 2,3,5,... remain prime, and, for them $R/(p)$ is $\mathbb{Z}/(p)$. But (nontrivial fact) if R is nonstandard it has nonstandard primes p, and for these $R/(p)$ is an infinite field. What properties does it have? (Answer later.)

Why are there infinite primes, if R nonstandard?

Overspill Principle (R nonstandard)

If X is first-order definable, and contains arbitrarily large standard elements, it contains a nonstandard element.

Proof Let δ be infinite positive, and let $Y = \{t < \delta : (\exists x \in X)(t < x)\}$. Note $\mathbb{N} \subseteq Y$. If result false, $\mathbb{N} = Y$. But $0 \in \mathbb{N}$, and closed under $u \mapsto u + 1$. So $\mathbb{N} = N$, contradiction. $\square$

Cor: There are infinite primes.

Proof Euclid. $\square$

A more delicate question is: If $x \in N$, $\exists y > x, y$ prime? (i.e. are primes cofinal).

Euclid's proof for $\mathbb{N}$ uses factorial. What meaning in N?

3 Exponentiation. Gödel's Device

In a general DOR, 2^x has no sense for $x \in N$. But Gödel showed that there is a natural definition in PA. The key is coding sequences.

Gödel's idea was to code arbitrary sequences $(a_0, ..., a_k)$ in terms of k and two auxiliary parameters λ, μ.

Define $\mathrm{Rem}(m, n)$, for $m, n \in N$ as the (clearly unique) ℓ so that $m = k.n + \ell$ in Euclidean division $(0 \leq \ell < n)$.

Given λ, μ, k we get a sequence $(a_0, ...a_k)$ as

$$a_j = \mathrm{Rem}(\lambda, 1 + (j+1).\mu) \qquad 0 \leq j \leq k.$$

The main point is that by suitable choice of λ, μ we can get **any** $(a_0, ...a_k)$. For now we work in $\mathbb{N}$, and use Chinese Remainder Theorem, which says:

a) if $n_0 ....n_k$ are coprime in pairs, and $0 \le a_i < n_i \quad 0 \le i \le k$ then there is a λ with $\mathrm{Rem}(\lambda, n_i) = a_i$;

b) moreover, λ can be chosen with $0 \le \lambda \le \prod n_k$.

Now, suppose p is a prime dividing $1 + (j+1).\mu$ and $1 + (\ell + 1).\mu \quad 0 \le j, \ell \le k$, then $p \mid (j - \ell).\mu$. So $p \mid (j - \ell)$ or $p \mid \mu$. But $p + \mu$ for then $p \mid 1$. So $p \mid (j - \ell)$. If $k! \mid \mu$ this implies that $p \mid \mu$, again a contradiction.

So if $k! \mid \mu$ the $i + (j+1).\mu \quad (0 \le j \le k)$ are coprime in pairs, and can be taken as the n_i above provided $a_j \le 1 + (j+1).\mu$.

Let $A = \max a_j$, and assume $\mu > A$. Then there is λ so that $\mathrm{Rem}(\lambda, 1 + (j+1).\mu) = a_j \quad 0 \le j \le k$, and $\lambda \le \pi(1 + (j+1).\mu)$. So, given $(a_0, ...a_k)$ we can choose $\mu \le \max(k!, \max a_j)$, and $\lambda \le \pi(1 + (j+1).\mu) \le (k+z)!.\mu^{k+1}$. These are rather large bounds, and later we try to bypass them.

Definition of the relation $y = 2^x$

Idea: Code $(2^0, 2^1, ...2^x)$

$$
\begin{aligned}
y &= 2^x \Leftrightarrow \exists \lambda, \mu[\mathrm{Rem}(\lambda, \mu) = 1 \\
\wedge(\forall_j &< x)(\mathrm{Rem}(\lambda, 1 + (j+2).\mu) \\
&= 2\,\mathrm{Rem}(\lambda, 1 + (j+1).\mu)) \\
\wedge\,\mathrm{Rem}(\lambda, 1 + (x+1).\mu) &= y].
\end{aligned}
$$

This is clearly correct in $\mathbb{N}$.

What does the RHS formula $\Phi(x, y)$ give in R, for R a general model of PA?

Theorem 1 For each x in N there is a unique y in N with $\Phi(x, y)$ (i.e. Φ is graph of a function).

Proof Clearly $\exists ! y \Phi(0, y)$.

Suppose $\exists ! y \Phi(x, y)$.

Suppose first $\Phi(x+1, t)$ and let λ', μ' witness this. By truncating the Rem conditions to $j < x$ we see that $\Phi(x, t/2)$, so $t/2$ unique, so t unique.

The existence is trickier, and needs some

Digressions

i) gcd and lcm. Given m, n we got $km + \ell n$ dividing m, n, and this gives $\gcd(m,n)$, the greatest common divisor of m, n. Then let $\operatorname{lcm}(m,n) = \dfrac{mn}{\gcd(m,n)}$. Obviously $m, n \mid \operatorname{lcm}(m,n)$. Suppose $m, n \mid a$. We have

$$\frac{a}{m}.(km + \ell n) = ak + \ell n.\frac{a}{m}$$

so $n \mid \dfrac{a}{m}.(km + \ell n)$, i.e. $\dfrac{mn}{(km + \ell n)} \mid a$. Thus we have a genuine lcm.

ii) $\operatorname{lcm}(1, ..., n)$. This exists in PA, and is easily proved by induction. Suppose $\operatorname{lcm}(1, ..., n)$ exists, $= \alpha$ say. Then $\operatorname{lcm}(\alpha, n+1)$ is easily seen to be $\operatorname{lcm}(1, ..., n+1)$.

iii) Chinese Remainder Theorem

This says that the congruence system

$$\begin{aligned} x &\equiv \alpha \mod m \\ x &\equiv \beta \mod n \end{aligned}$$

has a solution iff and only if $\alpha - \beta \equiv 0 \mod \gcd(m,n)$, (and x can be chosen with $0 \leq x < \operatorname{lcm}(m,n)$.

It holds in PA. The necessity is obvious. Clearly by Euclidean division wlog $0 \leq \alpha < m$, $0 \leq \beta < n$.

$$\begin{aligned} \alpha - \beta &= \frac{\alpha - \beta}{\gcd(m,n)}.\gcd(m,n) \\ &= \frac{\alpha - \beta}{\gcd(m,n)}.(\lambda m + \mu n) \\ &= am + bn, \end{aligned}$$

so $\alpha - am = \beta + bn$ can be taken to solve the system.

Now use induction to get least nonnegative solution of the congruences, x_0 say. If $x_0 \geq \operatorname{lcm}(m,n)$, write $x_0 = k.\operatorname{lcm}(m,n) + \ell$, $0 \leq \ell < \operatorname{lcm}(m,n)$, so $\ell \equiv x \mod m$ and $\mod n$, so ℓ is solution too.

iv) A maximum principle

Suppose that a subset I of N is bounded above and definable [NONEXAMPLE, $\mathbb{N}$, if R nonstandard.] Then I has a greatest element. For otherwise

$$J = \{n \in I : (\exists m \in I) n < m\} \cup \{0\}$$

violates induction.

v) **General lcm**

If I is as in (iv), there is an element (written $\operatorname{lcm} I$) which is divisible by all $n \in I$, and divides all other elements divisible by all elements of I.

Proof Let α be upper bound. Consider $\delta = \operatorname{lcm}(1, ..., \alpha)$ as before, and let J be the set of all numbers $\leq \delta$ which are divisible by all elements of I. Let $\delta_1 = \sup I$ (from (iv)). If $n \in I \rightarrow n \mid \delta_2$, let $\delta_2 = a\delta_1 + b \quad 0 \leq b < \delta_1$, to get $b = 0$ by minimality.

Cor: The primes are cofinal in N.

Proof Suppose α is a bound for the primes. Let $I =$ set of primes $(\leq \alpha)$.

Let $\delta = \operatorname{lcm} I$.

Let $q \mid \delta + 1,\ q$ prime, etc.

vi) **General CRT**

Let I be a bounded set, as in (iv), and in addition distinct members of I are pairwise coprime. Let $f : I \rightarrow N$ have definable graph. Then there is $x \in N,\ 0 \leq x \leq \operatorname{lcm} I,\ x \equiv f(n) \mod n$, all $n \in I$.

Proof Show by induction on k the result for $I \cap \{y : y \leq k\}$. If result for k, with solution x' say, get to $k+1$ as follows. Nothing to prove if $k+1 \notin I$, use x'. If $k+1 \in I$, consider $\operatorname{lcm} I \cap \{y : y \leq k\}$. Again show by induction this is coprime to $k+1$, so use CRT to get x'' with $x'' \equiv x \mod \operatorname{lcm} I \cap \{y : y \leq k\},\ x'' \equiv f(k+1) \mod k+1$, and we are done.

END OF DIGRESSION!

We have $\Phi(x, y)$ from Gödel's definition, witnessed by λ, μ. Let $A = 1 + \max(2y, 1+(x+1).\mu)$. Let $\mu' = \operatorname{lcm}(1, 2, ...(x+1)).A$. Let $I = \{1+(j+1).\mu' : 0 \leq j \leq (x+1)\}$ (which is bounded definable). The elements of I are pairwise coprime. So by general CRT get

$$\begin{aligned} \lambda' &= \operatorname{Rem}(\lambda, 1 + (j+1).\mu) \mod (1 + (j+1).\mu') \quad 0 \leq j \leq x \\ &= 2y \mod (1 + (x+2).\mu'). \end{aligned}$$

One easily checks $\Phi(x+1, 2y)$. This proves Theorem 1.

Additional information

Define 2^x (uniquely) by $\Phi(x, 2^x)$. We have from above $2^{x+1} = 2.2^x$. But in fact:

a) $2^{x_1+x_2} = 2^{x_1}.2^{x_2}$.

Proof Induction on x_2. □

b) $2^x > x$.

Proof Easy induction.

NOTE: We have not yet defined 3^x etc., but this is easy once has the CRT.

"Uniqueness" of 2^x

Theorem 2 There is a unique function $f : N \to N$ **with definable graph** so that

$$f(0) = 1,\ f(x+1) = 2.f(x).$$

Proof Take two solutions and prove by induction they agree. □

This is what we mean by 2^x in PA.

NOTE: In any countable nonstandard model of PA, there are undefinable solutions f, g, h of

$$\begin{aligned} f(0) &= 1 \\ f(x+1) &= f(x)+1 \\ g(0) &= 0 \\ g(x+1) &= g(x)+2 \\ h(0) &= 1 \\ h(x+1) &= h(x).2 \end{aligned}$$

so that

(i) $f(x) \neq x+1$ in general

(ii) $g(x) \neq 2x$ in general

(iii) $h(x) \neq 2^x$ in general.