domain of every elementarily equivalent subinterpretation $\mathcal{J}$. Each such $\mathcal{J}$ will therefore have a non-enumerable domain.

13.3 Suppose X is a nonempty set, and for any x in X there is a y in X such that xRy. Let $\mathcal{I}$ be the interpretation whose domain is X, and according to which '$\mathbf{R}$' is true of x, y iff xRy. $\forall x \exists y\, x\mathbf{R}y$ is true in $\mathcal{I}$. Let $\mathcal{J}$ be an elementarily equivalent subinterpretation of $\mathcal{I}$ whose domain E is enumerable. Let $e_0, e_1, e_2, \ldots$ be an enumeration of E. $\forall x \exists y\, x\mathbf{R}y$ is true in $\mathcal{J}$. Therefore for every e_i in E there will be an e_j in E such that $e_i R e_j$. Define f by: $f(0) = e_0$; for each n, $f(n+1) = e_j$ iff $f(n)\ Re_j$ and for every $k < j$, not: $f(n)\ Re_k$. The axiom of dependent choice is not required to guarantee the existence of f.

14
Representability in Q

The present chapter falls into three parts. In the first part we introduce the notion of *representability of a function* (of natural numbers) *in a theory* and present a theory, called 'Q'. In the second part we give an alternative characterization of the recursive functions,† and in the third we use this new characterization to show that every recursive function is representable in the theory Q. In the next chapter several important results about undecidability, indefinability and incompleteness will be shown to follow from the latter result. The converse, that every function representable in Q is recursive, is also true, and we shall also indicate why at the end of the next chapter (Exercise 15.2).

Part I

We shall take a theory to be a set of sentences in some language that contains all of its logical consequences that are sentences in that language. If a sentence A is a member of theory T, it is called a *theorem* of T; to indicate that A is a theorem of T, we write: $\vdash_T A$.

From now through Chapter 21, we shall confine our attention to *numerical* theories: theories whose language contains the name $\mathbf{0}$ and the one-place function symbol $'$. (Q will be such a theory.) The *numeral* for n, $\mathbf{n}$, is the result of attaching n occurrences of $'$ to (the right of) $\mathbf{0}$. Thus $\mathbf{3} = \mathbf{0}'''$ and the numeral for $n+1$ is $\mathbf{n}'$. For any natural number n, $\mathbf{n}$ is an expression or sequence of symbols, a *term* of the sort described.

If A is a formula that contains free occurrences of the n (distinct) variables $x_1, \ldots, x_n$, we shall sometimes refer to A as $A(x_1, \ldots, x_n)$. For any natural numbers $p_1, \ldots, p_n$, $A(\mathbf{p}_1, \ldots, \mathbf{p}_n)$ is the result of substituting an occurrence of $\mathbf{p}_i$ for each free occurrence of x_i in $A(x_1, \ldots, x_n)$ (for each i between 1 and n). In discussing a formula $A(x_1, \ldots, x_n)$ we may wish to consider a formula, which we refer to as '$A(y_1, \ldots, y_n)$'. This is to be understood to be the formula that results when any bound occurrence of y_i that may occur in $A(x_1, \ldots, x_n)$ is first replaced by an occurrence of a new

† Cf. the last paragraph of Chapter 8.

variable z_i (different z_is for different y_is), and then an occurrence of y_i is substituted for each free occurrence of x_i in the result.

To reduce clutter, we shall write '$\mathfrak{p}$' instead of '$\mathfrak{p}_1, \ldots, \mathfrak{p}_n$', '$\mathfrak{x}$' instead of '$x_1, \ldots, x_n$', and '$\mathfrak{P}$' instead of '$\mathfrak{p}_1, \ldots, \mathfrak{p}_n$'.

We can now define representability. An n-place function f is *representable* in a theory T if there is a formula $A(\mathfrak{x}, x_{n+1})$ such that for any natural numbers $\mathfrak{p}, j$, if $f(\mathfrak{p}) = j$, then $\vdash_T \forall x_{n+1}(A(\mathfrak{P}, x_{n+1}) \leftrightarrow x_{n+1} = \mathfrak{j})$. In this case the formula $A(\mathfrak{x}, x_{n+1})$ is said to *represent* f in T.

The requirement that $\vdash_T \forall x_{n+1}(A(\mathfrak{P}, x_{n+1}) \leftrightarrow x_{n+1} = \mathfrak{j})$ should hold whenever $f(\mathfrak{p}) = j$ is equivalent to the requirement that both $\vdash_T A(\mathfrak{P}, \mathfrak{j})$ and $\vdash_T \forall x_{n+1}(A(\mathfrak{P}, x_{n+1}) \rightarrow x_{n+1} = \mathfrak{j})$ should hold whenever $f(\mathfrak{p}) = j$. If the sentence $\mathfrak{j} \neq \mathfrak{k}$ is a theorem of T whenever $j \neq k$ (and we shall see that Q is a theory of which this is so), then if A represents f in T and $f(\mathfrak{p}) \neq k$, then $\vdash_T -A(\mathfrak{P}, \mathfrak{k})$ (for $\vdash_T \mathfrak{j} \neq \mathfrak{k}$, where $j = f(\mathfrak{p})$).

The language of theory Q is L, the *language of arithmetic*. L contains four non-logical symbols, the name $\mathfrak{o}$, the one-place function symbol $'$, and two two-place function symbols, $+$ and $\cdot$. Q is the set of sentences in L that are logical consequences of these seven sentences, the *axioms of* Q:

$Q1$ $\forall x \forall y(x' = y' \rightarrow x = y)$,

$Q2$ $\forall x\, \mathfrak{o} \neq x'$,

$Q3$ $\forall x(x \neq \mathfrak{o} \rightarrow \exists y\, x = y')$,

$Q4$ $\forall x\, x + \mathfrak{o} = x$,

$Q5$ $\forall x \forall y\, x + y' = (x + y)'$,

$Q6$ $\forall x\, x \cdot \mathfrak{o} = \mathfrak{o}$,

$Q7$ $\forall x \forall y\, x \cdot y' = (x \cdot y) + x$.

Q is a consistent theory, for all of its axioms are true in the *standard interpretation* $\mathcal{N}$ *for its language* L, in which the domain is the set of all natural numbers, $\mathfrak{o}$ is assigned zero as denotation, and $'$, $+$, and $\cdot$ are assigned the successor, addition, and multiplication functions. Q is a theory that is rather strong in certain ways (all recursive functions are representable in it), but rather weak in others (e.g. $\forall x \forall y\, x + y = y + x$ is not a theorem of Q, as an exercise at the end of the chapter shows). Tarski, Mostowski, and R. Robinson have written that it 'is distinguished by the simplicity and clear mathematical content of its axioms'. We shall devote the remainder of this chapter to showing that all recursive functions are representable in Q.

Part II

We recall from Chapters 7 and 8 that the recursive functions can be characterized as those functions obtainable from the zero function, the successor function and the identity functions by means of a finite number of applications of the operations of composition, primitive recursion, and minimization of those functions called *regular* functions.

The *zero* function z is the one-place function whose value for all arguments is zero.

The *successor* function $'$ $(= s)$ is the one-place function whose value for any argument i is $i + 1$ $(= i'$, the successor of $i)$.

For each $m \geqslant 1$ and each $n \leqslant m$, there is an m-place *identity* function id_n^m. For any natural numbers $i_1, \ldots, i_m$, $\mathrm{id}_n^m(i_1, \ldots, i_m) = i_n$.

If f is an m-place function, and $g_1, \ldots, g_m$ are all n-place functions, then the n-place function h is said to be obtained from $f, g_1, \ldots, g_m$ by *composition* if for any natural numbers $\mathfrak{p}$, $h(\mathfrak{p}) = f(g_1(\mathfrak{p}), \ldots, g_m(\mathfrak{p}))$.

If f is an n-place function and g is an $(n+2)$-place function, then the $(n+1)$-place function h is said to be obtained from f and g by *primitive recursion* if for any natural numbers $\mathfrak{p}$, k, $h(\mathfrak{p}, \mathfrak{o}) = f(\mathfrak{p})$ and

$$h(\mathfrak{p}, k+1) = g(\mathfrak{p}, k, h(\mathfrak{p}, k)).$$

An $(n+1)$-place function f is called *regular* if for any natural numbers $\mathfrak{p}$, there exists at least one natural number i such that $f(\mathfrak{p}, i) = \mathfrak{o}$. If f is a regular $(n+1)$-place function, then the n-place function g is said to be obtained from f by *minimization* if for any natural numbers $\mathfrak{p}$,

$$g(\mathfrak{p}) = \mu i f(\mathfrak{p}, i) = \mathfrak{o},$$

where 'μi' means 'the least natural number i such that'.

If R is an n-place relation of natural numbers (i.e. a set of ordered n-tuples of natural numbers), then the *characteristic function* of R is the n-place function f_R such that for any $\mathfrak{p}$,

$$f_R(\mathfrak{p}) = \begin{cases} 1 & \text{if } R\mathfrak{p} \text{ (i.e. if } \mathfrak{p} \text{ is in } R), \\ \mathfrak{o} & \text{if not } R\mathfrak{p}. \end{cases}$$

$f_=$ is thus the characteristic function of the identity relation. For any i, j, $f_=(i,j) = 1$ if $i = j$ and $f_=(i,j) = \mathfrak{o}$ if $i \neq j$.

We shall call a function *Recursive* (capital 'R') if it can be obtained from the functions $+$, $\cdot$, $f_=$, and the various id_n^m by means of a finite number of applications of the two operations of composition and minimization of regular functions.

All Recursive functions are recursive, for $+$, $\cdot$, $f_=$ and the functions id_n^m are recursive, and the recursive functions are closed under composition and minimization of regular functions. On the other hand the zero function z is Recursive, for, as we saw in Chapter 7, z can be obtained from $\cdot$ by minimization. And s is obtainable by composition from Recursive functions and thus is Recursive too: for all i,

$$s(i) = i+1 = \mathrm{id}_1^1(i)+f_=(\mathrm{id}_1^1(i), \mathrm{id}_1^1(i)).$$

In the rest of Part II we show that all other recursive functions are also Recursive, and for this it suffices to show that if f and g are Recursive functions from which h is obtained by primitive recursion, then h is also Recursive.

We must first see that certain relations and functions are Recursive. A relation is Recursive iff its characteristic function is Recursive. (So $=$ is Recursive.)

Suppose, for example, that d is a two-place Recursive function and e is a three-place Recursive function. Let R be the 6-place relation defined by: Ri,j,k,m,n,q iff $d(j,n) = e(n,k,m)$. Then R is Recursive, for

$$f_R(i,j,k,m,n,q) = f_=(d(\mathrm{id}_2^6(i,j,k,m,n,q), \mathrm{id}_5^6(i,j,k,m,n,q)),$$
$$e(\mathrm{id}_5^6(i,j,k,m,n,q), \mathrm{id}_3^6(i,j,k,m,n,q), \mathrm{id}_4^6(i,j,k,m,n,q))).$$

Similarly, all other relations obtained by 'setting Recursive functions equal to each other' are Recursive.

Suppose that R and S are n-place Recursive relations. Then the intersection $(R\&S)$ of R and S and the complement $-R$ of R are Recursive, for

$$f_{(R\,\&\,S)}(\mathsf{p}) = f_R(\mathsf{p})\cdot f_S(\mathsf{p}), \text{ and } f_{-R}(\mathsf{p}) = f_=(f_R(\mathsf{p}), z(f_R(\mathsf{p}))) \,(= f_=(f_R(\mathsf{p}), 0)).$$

As & and $-$ suffice to define all truth-functional connectives, any relation obtained from Recursive relations by truth-functional, i.e., Boolean, operations is also Recursive. E.g. if Ri,j,k if and only if either $i = k$ or $k \neq j$, then R is Recursive.

If R is an $(n+1)$-place relation, then e will be said to be obtained from R by *minimization* if for any p, $e(\mathsf{p}) = \mu i R\mathsf{p}, i$. ($e$ may be undefined for some p.) An $(n+1)$-place relation will be called *regular* if for any p, there is an i such that $R\mathsf{p}, i$. The function obtained from a regular relation by minimization is everywhere defined.

If R is a regular, Recursive $(n+1)$-place relation, and e is obtained from R by minimization, then e is Recursive, for $e(\mathsf{p}) = \mu i f_{-R}(\mathsf{p}, i) = 0$. ($R\mathsf{p}, i$ iff $f_{-R}(\mathsf{p}, i) = 0$.)

Finally, if R is an $(n+1)$-place relation, then the $(n+1)$-place relation S will be said to be obtained from R by *bounded universal quantification*† if (for all p, j) $S\mathsf{p}, j$ iff $\forall i < j\, R\mathsf{p}, i$. If R is Recursive and S is obtained from R by bounded universal quantification, then S is Recursive. *Proof.* Let T be defined by: $T\mathsf{p}, j, i$ iff either not $R\mathsf{p}, i$ or $i = j$. T is regular (for all p, j, $T\mathsf{p}, j, j$) and Recursive (by the foregoing). Let d be defined by: $d(\mathsf{p}, j) = \mu i T\mathsf{p}, j, i$. d is Recursive. For any p, j, $d(\mathsf{p}, j) \leqslant j$. And $d(\mathsf{p}, j) = j$ iff for every $i < j$, $R\mathsf{p}, i$; iff $S\mathsf{p}, j$. So if e is defined by: $e(\mathsf{p}, j) = f_=(j, d(\mathsf{p}, j))$, then e is Recursive and the characteristic function of S.

S is said to be obtained from R by *bounded existential quantification*† if (for all p, j) $S\mathsf{p}, j$ iff $\exists i < j\, R\mathsf{p}, i$. Analogously, any relation obtained from a Recursive relation by bounded existential quantification is Recursive.

We'll now define J, the pairing function.

Definition

$$J(a,b) = \tfrac{1}{2}(a+b)(a+b+1)+a.$$

Lemma 14.1

J is a one–one function whose domain is the set of all ordered pairs $\langle a, b\rangle$ of natural numbers and whose range is the set of all natural numbers.

Proof. There are $n+1$ pairs $\langle a, b\rangle$ such that $a+b = n$ (*viz.*, $\langle 0, n\rangle$, $\langle 1, n-1\rangle, \ldots, \langle n, 0\rangle$). So there are $0+1+2+\ldots+n$, $= \frac{1}{2}n(n+1)$, pairs $\langle c, d\rangle$ such that $c+d < n$. We'll say that $\langle c, d\rangle$ *precedes* $\langle a, b\rangle$ *in order* O (cf. Chapter 13) if either $c+d < a+b$ or ($c+d = a+b$ and $c < a$). There are a natural numbers less than a. So if $a+b = n$, there are $\frac{1}{2}n(n+1)+a$ pairs that precede $\langle a, b\rangle$ in order O. But if $a+b = n$, then

$$\tfrac{1}{2}n(n+1)+a = J(a,b).$$

So $J(a, b)$ is precisely the number of pairs preceding $\langle a, b\rangle$ in order O.

$a, b \leqslant J(a, b)$. J is Recursive, for J is obtained from a regular Recursive function by minimization: $J(a, b) = \mu i[i+i = (a+b)(a+b+1)+2a]$.

Define K and L, the inverse pairing functions, by:

$K(i) = \mu a \exists b \leqslant i\, J(a,b) = i$ (i.e. $\mu a[\exists b < i\, J(a,b) = i \vee J(a,i) = i]$),
$L(i) = \mu b \exists a \leqslant i\, J(a,b) = i$. By Lemma 14.1, K and L are Recursive.

† These definitions differ slightly from those given in Chapter 7 in that '$<$' is used instead of '$\leqslant$'.

We now define some more relations and functions; it should be evident from their definitions that they are Recursive.

m *divides* $n \leftrightarrow \exists i \leqslant n\, i \cdot m = n$.

p *is prime* $\leftrightarrow \{p \neq 0 \,\&\, p \neq 1 \,\&\, \forall m \leqslant p[m \text{ divides } p \rightarrow (m = 1 \vee m = p)]\}$.

$m < n \leftrightarrow \exists i < n\, i = m$.

$m \dot{-} n = \mu i([n < m \rightarrow n + i = m] \,\&\, [-n < m \rightarrow i = 0])$.

n *is a power of the prime* $p \leftrightarrow \{n \neq 0 \,\&\, p \text{ is prime} \,\&\, \forall m \leqslant n[m \text{ divides } n \rightarrow (m = 1 \vee p \text{ divides } m)]\}$.

(Notice that we can't simply say that n is a power of k iff for every $m \leqslant n$, if m divides n, then $m = 1$ or k divides m; let $n = k = 6$, $m = 2$.)

$\eta(p, b) = \mu i[(p \text{ is prime} \,\&\, i \text{ is a power of the prime } p \,\&\, i > b \,\&\, i > 1) \vee (p \text{ is not prime} \,\&\, i = 0)]$.

For prime p, $\eta(p, b)$ is the least number whose base p numeral is *longer* than the base p numeral for b. E.g. $\eta(7, 25) = 49$. (Note that $25 = 34_7$ and $49 = 100_7$.)

$a *_p b = a \cdot \eta(p, b) + b$.

If $a \neq 0$, $a *_p b$ is $\neq 0$ and is the number denoted in base p notation by the result of writing the base p numeral for b directly to the right of that for a. So, e.g. $4 *_7 25 = 4 \cdot 49 + 25 = 221$, and $4_7 = 4$, $34_7 = 25$, and $434_7 = 221$. In what follows, association is assumed to be to the left: '$a *_p b *_p c$' means '$(a *_p b) *_p c$', not '$a *_p (b *_p c)$'. Then if $a \neq 0$, $a *_p b *_p c *_p \ldots *_p z$ is the number denoted in base p notation by the result of writing down the base p numeral for b directly to the right of that for a, then that for c directly to the right of *that*, ... and then that for z directly to the right of *that*.

$a \text{ part}_p b \leftrightarrow \exists c \leqslant b \exists d \leqslant b\, [c *_p a *_p d = b \vee c *_p a = b \vee a *_p d = b \vee a = b]$.

$a \text{ part}_p b$ iff $a = 0$ or $a = b$ or b's base p numeral can be obtained by attaching base p numerals to the left and/or right of a's base p numeral.

$\alpha(p, q, j) = \mu i[(p \dot{-} 1) *_p j *_p i \text{ part}_p q \vee i = q]$. ('$i = q$' is for 'waste cases'.)

$\beta(i, j) = \alpha(K(i), L(i), j)$.

Lemma 14.2. (*The β-function lemma*)
For any k and any finite sequence of natural numbers $i_0, \ldots, i_k$, there exists a natural number i such that for every $j \leqslant k$, $\beta(i, j) = i_j$.

Proof. Let $i_0, \ldots, i_k$ be a finite sequence of natural numbers. Let p be a prime such that $p - 1$ is greater than all of $i_0, \ldots, i_k, k$. (There are infinitely many primes.) Let $s = p - 1$. $s \neq 0$. All of $s, 0, i_0, \ldots, k, i_k$ are represented by single digits in base p notation (!). Let

$$q = s *_p 0 *_p i_0 *_p s *_p 1 *_p i_1 *_p \ldots *_p s *_p k *_p i_k.$$

Then for every $j \leqslant k$, $\alpha(p, q, j) = i_j$. Let $i = J(p, q)$. Then for every $j \leqslant k$, $\beta(i, j) = i_j$.

Suppose now that f is an n-place function, that g is an $(n+2)$-place function, and that h is obtained from f and g by primitive recursion. Then $h(\mathsf{p}, 0) = f(\mathsf{p})$ and (for any k) for every $j < k$, $h(\mathsf{p}, j') = g(\mathsf{p}, j, h(\mathsf{p}, j))$. By the β-function lemma, for any k there is an i such that for every $j \leqslant k$, $\beta(i, j) = h(\mathsf{p}, j)$. These is are precisely those such that $\beta(i, 0) = f(\mathsf{p})$ and for every $j < k$, $\beta(i, j') = g(\mathsf{p}, j, \beta(i, j))$. Therefore, if R is the $(n+2)$-place relation defined by:

$$R\mathsf{p}, k, i \text{ iff } \beta(i, 0) = f(\mathsf{p}) \,\&\, \forall j < k\, \beta(i, j') = g(\mathsf{p}, j, \beta(i, j)),$$

then R is regular; and R is Recursive if f and g are. So if d is the $(n+1)$-place function defined by: $d(\mathsf{p}, k) = \mu i R\mathsf{p}, k, i$, then d is Recursive if f and g are. Moreover $d(\mathsf{p}, k)$ is the least i such that for every $j \leqslant k$, $\beta(i, j) = h(\mathsf{p}, j)$. For any such i, $\beta(i, k) = h(\mathsf{p}, k)$. We may thus define h by composition from β, d, id^{n+1}_{n+1}: $h(\mathsf{p}, k) = \beta(d(\mathsf{p}, k), \mathrm{id}^{n+1}_{n+1}(\mathsf{p}, k))$. As β and id^{n+1}_{n+1} are Recursive, h is Recursive if f and g are Recursive.

Thus any function obtained by primitive recursion from Recursive functions is itself Recursive.

We have therefore shown that a function is recursive if and only if it is Recursive.

Part III

We'll now show that all Recursive functions are representable in Q, from which we conclude that all recursive functions are representable in Q.

The identity functions id^m_n are all representable in Q: since for any $i_1, \ldots, i_m$, $\forall x_{m+1}((\mathsf{i}_1 = \mathsf{i}_1 \,\&\, \ldots \,\&\, \mathsf{i}_m = \mathsf{i}_m \,\&\, x_{m+1} = \mathsf{i}_n) \leftrightarrow x_{m+1} = \mathsf{i}_n)$ is *valid*,

$$(x_1 = x_1 \,\&\, \ldots \,\&\, x_m = x_m \,\&\, x_{m+1} = x_n)$$

represents id^m_n in Q.

We now show that addition is represented in Q by the formula

$$x_1 + x_2 = x_3.$$

Lemma 14.3

Suppose that $i+j = k$. Then $\vdash_Q \mathbf{i}+\mathbf{j} = \mathbf{k}$.

Proof. The proof is an induction on j. Basis step: $j = 0$. We must show that $\vdash_Q \mathbf{i}+\mathbf{0} = \mathbf{i}$. But this follows from $Q4$. Induction step: $j = m'$. Then for some n, $k = n'$ and $i+m = n$, whence by the induction hypothesis, $\vdash_Q \mathbf{i}+\mathbf{m} = \mathbf{n}$, and therefore $\vdash_Q (\mathbf{i}+\mathbf{m})' = \mathbf{n}'$. Since

$$\vdash_Q (\mathbf{i}+\mathbf{m})' = \mathbf{i}+\mathbf{m}'$$

by $Q5$, it follows that $\vdash_Q \mathbf{i}+\mathbf{j} = \mathbf{k}$.

Lemma 14.4

$x_1+x_2 = x_3$ represents addition in Q.

Proof. $\forall x_3(\mathbf{i}+\mathbf{j} = x_3 \leftrightarrow x_3 = \mathbf{k})$ is a logical consequence of $\mathbf{i}+\mathbf{j} = \mathbf{k}$, which, by 14.3, is a theorem of Q if $i+j = k$.

Multiplication:

Lemma 14.5

Suppose that $i \cdot j = k$. Then $\vdash_Q \mathbf{i} \cdot \mathbf{j} = \mathbf{k}$.

Proof. Induction on j. If $j = 0$, we must show that $\vdash_Q \mathbf{i} \cdot \mathbf{0} = \mathbf{0}$. But this follows from $Q6$. If $j = m'$, then $k = n+i$, where $n = i \cdot m$. By the hypothesis of the induction, $\vdash_Q \mathbf{i} \cdot \mathbf{m} = \mathbf{n}$. By 14.3, $\vdash_Q \mathbf{n}+\mathbf{i} = \mathbf{k}$. By $Q7$, $\vdash_Q \mathbf{i} \cdot \mathbf{m}' = \mathbf{i} \cdot \mathbf{m}+\mathbf{i}$. So $\vdash_Q \mathbf{i} \cdot \mathbf{m}' = \mathbf{k}$, i.e., $\vdash_Q \mathbf{i} \cdot \mathbf{j} = \mathbf{k}$.

Lemma 14.6

$x_1 \cdot x_2 = x_3$ represents multiplication in Q.

Proof. This follows from 14.5 just as 14.4 followed from 14.3.

So $+$ and $\cdot$ are representable in Q.

Let's now verify that if $i \neq j$, then $\mathbf{i} \neq \mathbf{j}$ is a theorem of Q.

Lemma 14.7

If $i \neq j$, then $\vdash_Q \mathbf{i} \neq \mathbf{j}$.

Proof. We may suppose without loss of generality that $i < j$. Induction on i. If $i = 0$, then $j > 0$, and so for some n, $j = n'$. We must show that $\vdash_Q \mathbf{0} \neq \mathbf{j}$, i.e., that $\vdash_Q \mathbf{0} \neq \mathbf{n}'$. But this immediately follows from $Q2$. If $i = m'$, then $j = n'$ and $m < n$, for some n. By the induction hypothesis, $\vdash_Q \mathbf{m} \neq \mathbf{n}$, and hence by $Q1$, $\vdash_Q \mathbf{m}' \neq \mathbf{n}'$, i.e., $\vdash_Q \mathbf{i} \neq \mathbf{j}$.

Lemma 14.8

Let $A(x_1, x_2, x_3)$ = the formula

$$(x_1 = x_2 \,\&\, x_3 = \mathbf{1}) \vee (x_1 \neq x_2 \,\&\, x_3 = \mathbf{0}).$$

Then $A(x_1, x_2, x_3)$ represents $f_=$ in Q.

Proof. If $f_=(i,j) = 1$, then $i = j$. So $\vdash_Q \mathbf{i} = \mathbf{j} \,\&\, \mathbf{1} = \mathbf{1}$, so $\vdash_Q A(\mathbf{i}, \mathbf{j}, \mathbf{1})$, whence $\vdash_Q \forall x_3(A(\mathbf{i}, \mathbf{j}, x_3) \leftrightarrow x_3 = \mathbf{1})$, as $\forall x_3 (A(\mathbf{i}, \mathbf{j}, x_3) \rightarrow x_3 = \mathbf{1})$ is a logical consequence of $A(\mathbf{i}, \mathbf{j}, \mathbf{1})$ when $i=j$. If $f_=(i,j)=0$, then $i \neq j$. By 14.7 $\vdash_Q \mathbf{i} \neq \mathbf{j}$. So $\vdash_Q \mathbf{i} \neq \mathbf{j} \,\&\, \mathbf{0} = \mathbf{0}$, whence $\vdash_Q \forall x_3(A(\mathbf{i}, \mathbf{j}, x_3) \leftrightarrow x_3 = \mathbf{0})$.

Thus $f_=$ is also representable in Q. We now show that any function obtained by composition from functions representable in Q is also representable in Q.

Suppose that $A(x_1, \ldots, x_m, x)$ represents f in Q, and that

$$B_1(\mathbf{x}, x_{n+1}), \ldots, B_m(\mathbf{x}, x_{n+1})$$

represent $g_1, \ldots, g_m$, respectively. Then if h is obtained from $f, g_1, \ldots, g_m$, by composition,

$$C(\mathbf{x}, x), = \exists y_1 \ldots \exists y_m(B_1(\mathbf{x}, y_1) \,\&\, \ldots \,\&\, B_m(\mathbf{x}, y_m) \,\&\, A(y_1, \ldots, y_m, x)),$$

represents h.

For if $g_1(\mathbf{p}) = i_1, \ldots, g_m(\mathbf{p}) = i_m$, and $f(i_1, \ldots, i_m) = j$, then $h(\mathbf{p}) = j$, and

$$\vdash_Q B_1(\mathbf{p}, \mathbf{i}_1), \tag{1}$$

$$\vdash_Q \forall x_{n+1}(B_1(\mathbf{p}, x_{n+1}) \rightarrow x_{n+1} = \mathbf{i}_1), \tag{2}$$

$$\vdots$$

$$\vdash_Q B_m(\mathbf{p}, \mathbf{i}_m), \tag{2m-1}$$

$$\vdash_Q \forall x_{n+1}(B_m(\mathbf{p}, x_{n+1}) \rightarrow x_{n+1} = \mathbf{i}_m), \tag{2m}$$

$$\vdash_Q A(\mathbf{i}_1, \ldots, \mathbf{i}_m, \mathbf{j}), \text{ and} \tag{2m+1}$$

$$\vdash_Q \forall x(A(\mathbf{i}_1, \ldots, \mathbf{i}_m, x) \rightarrow x = \mathbf{j}). \tag{2m+2}$$

(1), (3), ..., $(2m-1)$, and $(2m+1)$ clearly entail that $\vdash_Q C(\mathbf{p}, \mathbf{j})$. And (2), (4), ..., $(2m)$, and $(2m+2)$ entail that $\vdash_Q \forall x(C(\mathbf{p}, x) \to x = \mathbf{j})$. We may see this as follows: Assume we have $B_1(\mathbf{p}, y_1), \ldots, B_m(\mathbf{p}, y_m)$, and $A(y_1, \ldots, y_m, x)$. From (2), we have $y_1 = \mathbf{i}_1, \ldots$, and from $(2m)$ we have $y_m = \mathbf{i}_m$. So we have $A(\mathbf{i}_1, \ldots, \mathbf{i}_m, x)$, whence from $(2m+2)$ we have $x = \mathbf{j}$. Thus

$$\vdash_Q \forall x(\exists y_1, \ldots \exists y_m(B_1(\mathbf{p}, y_1) \,\&\, \ldots \,\&\, B_m(\mathbf{p}, y_m) \,\&\, A(y_1, \ldots, y_m, x)) \to x = \mathbf{j}),$$

i.e. $\vdash_Q \forall x(C(\mathbf{p}, x) \to x = \mathbf{j})$, and therefore C represents h.

Lemma 14.9

For each i, $\vdash_Q \forall x\, x' + \mathbf{i} = x + \mathbf{i}'$.

Proof. Induction on i. If $i = 0$, $\forall x\, x' + \mathbf{0} = x + \mathbf{0}'$ follows from

$$\forall x(x' + \mathbf{0} = x' = (x + 0)' = x + \mathbf{0}'),$$

which follows from Q_4 and Q_5. If $i = m'$, then by the induction hypothesis $\vdash_Q \forall x\, x' + \mathbf{m} = x + \mathbf{m}'$, whence by Q_5, $\vdash_Q \forall x(x' + \mathbf{m}' = (x' + \mathbf{m})' = (x + \mathbf{m}')' = x + \mathbf{m}'')$, and hence

$$\vdash_Q \forall x\, x' + \mathbf{i} = x + \mathbf{i}'.$$

We now define $x_1 < x_2$ to be the formula $\exists x_3\, x_3' + x_1 = x_2$.

Lemma 14.10

If $i < j$, then $\vdash_Q \mathbf{i} < \mathbf{j}$.

Proof. Suppose $i < j$. Then for some m, $m' + i = j$. By 14.3,

$$\vdash_Q \mathbf{m}' + \mathbf{i} = \mathbf{j}, \quad \text{and so} \quad \vdash_Q \exists x_3 x_3' + \mathbf{i} = \mathbf{j}, \text{ i.e. } \vdash_Q \mathbf{i} < \mathbf{j}.$$

Lemma 14.11

For each i, $\vdash_Q \forall x(x < \mathbf{i} \to x = \mathbf{0} \vee \ldots \vee x = \mathbf{i} - 1)$ (where, if $i = 0$, the consequent is an empty disjunction and hence is to be regarded as equivalent to $\mathbf{0} \neq \mathbf{0}$).

Proof. Induction on i. Basis step: $i = 0$. We must show $\vdash_Q \forall x\, \neg x < \mathbf{0}$. By Q_3 we have $x = \mathbf{0} \vee \exists y\, x = y'$. Assume $x < \mathbf{0}$, i.e., $\exists w\, w' + x = \mathbf{0}$. If $x = \mathbf{0}$ holds, we have $w' = w' + \mathbf{0}$ (by Q_4) $= w' + x = \mathbf{0}$, which is impossible by Q_2. If $x = y'$ holds, we have $(w' + y)' = w' + y'$ (by Q_5) $= w' + x = \mathbf{0}$ which is again impossible by Q_2. Thus $\vdash_Q \forall x\, \neg x < \mathbf{0}$.

Induction step. We suppose $\vdash_Q \forall x(x < \mathbf{i} \to x = \mathbf{0} \vee \ldots \vee x = \mathbf{i} - 1)$. We must show $\vdash_Q \forall x(x < \mathbf{i}' \to x = \mathbf{0} \vee x = \mathbf{0}' \vee \ldots \vee x = \mathbf{i})$. Assume we have

$x < \mathbf{i}'$, i.e., $\exists w\, w' + x = \mathbf{i}'$. By Q_3 we have $\exists y\, x = y' \vee x = \mathbf{0}$. If $x = y'$ holds, then we have $\mathbf{i}' = w' + x = w' + y' = (w' + y)'$ (by Q_5), whence by Q_1 we have $\mathbf{i} = w' + y$, and therefore $y < \mathbf{i}$. By the induction hypothesis we have $y = \mathbf{0} \vee \ldots \vee y = \mathbf{i} - 1$ (if $i = 0$, we have $\mathbf{0} \neq \mathbf{0}$), and therefore we have $x = \mathbf{0}' \vee \ldots \vee x = \mathbf{i}$ (if $i = 0$, we have $\mathbf{0} \neq \mathbf{0}$), and therefore we have $x = \mathbf{0} \vee x = \mathbf{0}' \vee \ldots \vee x = \mathbf{i}$, which we also have in case $x = \mathbf{0}$ holds.

Lemma 14.12

For each i, $\vdash_Q \forall x(\mathbf{i} < x \to x = \mathbf{i}' \vee \mathbf{i}' < x)$.

Proof. Assume $\mathbf{i} < x$, i.e. $\exists w w' + \mathbf{i} = x$. We have $w = \mathbf{0} \vee \exists y\, w = y'$ by Q_3. From $w = \mathbf{0}$ and $w' + \mathbf{i} = x$, we have $\mathbf{0}' + \mathbf{i} = x$, whence by 14.3 we have $x = \mathbf{i}'$. From $w = y'$ and $w' + \mathbf{i} = x$, we have $y'' + \mathbf{i} = x$, whence by 14.9 we have $y' + \mathbf{i}' = x$, and so we have $\mathbf{i}' < x$.

Lemma 14.13

For each i, $\vdash_Q \forall x(\mathbf{i} < x \vee x = \mathbf{i} \vee x < \mathbf{i})$.

Proof. Induction on i. Basis step: $i = 0$. Assume $x \neq \mathbf{0}$. By Q_3 we then have $\exists y\, x = y'$, and so by Q_4 we have $\exists y\, y' + \mathbf{0} = x$, i.e., $\mathbf{0} < x$. Induction step: we suppose $\vdash_Q \forall x(\mathbf{i} < x \vee x = \mathbf{i} \vee x < \mathbf{i})$. We must show $\vdash_Q \forall x(\mathbf{i}' < x \vee x = \mathbf{i}' \vee x < \mathbf{i}')$. By 14.12

$$\vdash_Q \forall x(\mathbf{i} < x \to x = \mathbf{i}' \vee \mathbf{i}' < x). \text{ By 14.10} \qquad \text{(I)}$$

$$\vdash_Q \forall x(x = \mathbf{i} \to x < \mathbf{i}'). \text{ And by 14.11 and 14.10} \qquad \text{(II)}$$

$$\vdash_Q \forall x(x < \mathbf{i} \to x < \mathbf{i}'). \qquad \text{(III)}$$

But from (I), (II), (III), and the induction hypothesis, it follows that

$$\vdash_Q \forall x(\mathbf{i}' < x \vee x = \mathbf{i}' \vee x < \mathbf{i}').$$

We can now show that the result g of applying minimization to any regular $(n+1)$-place function f that is representable in Q is also representable in Q.

Suppose that f is a regular $(n+1)$-place function, and that

$$A(\mathbf{x}, x_{n+1}, x_{n+2})$$

represents f in Q. Let $B(\mathbf{x}, x_{n+1}) =$ the formula

$$(A(\mathbf{x}, x_{n+1}, \mathbf{0}) \,\&\, \forall w\,(w < x_{n+1} \to \neg A(\mathbf{x}, w, \mathbf{0}))).$$

Then B represents g in Q.

For suppose that $g(\mathbf{p}) = i$. Then $f(\mathbf{p}, i) = \mathbf{0}$, and for any $j < i$, $f(\mathbf{p}, j) \neq \mathbf{0}$. Since A represents f in Q, we have

$$\vdash_Q A(\mathbf{p}, \mathbf{i}, \mathbf{0}), \text{ and (if } i > 0), \tag{i}$$

$$\vdash_Q \neg A(\mathbf{p}, \mathbf{0}, \mathbf{0}), \tag{0}$$

$$\vdots \qquad\qquad \vdots$$

$$\vdash_Q \neg A(\mathbf{p}, \mathbf{i}-\mathbf{1}, \mathbf{0}). \tag{$i-1$}$$

(0), ..., $(i-1)$, and 14.11 entail that

$$\vdash_Q \forall w(w < \mathbf{i} \rightarrow \neg A(\mathbf{p}, w, \mathbf{0})), \tag{$i+1$}$$

which, together with (i), entails that $\vdash_Q B(\mathbf{p}, \mathbf{i})$.

We must show that $\vdash_Q \forall x_{n+1}(B(\mathbf{p}, x_{n+1}) \rightarrow x_{n+1} = \mathbf{i})$. Assume $B(\mathbf{p}, x_{n+1})$, i.e., $A(\mathbf{p}, x_{n+1}, \mathbf{0}) \,\&\, \forall w(w < x_{n+1} \rightarrow \neg A(\mathbf{p}, w, \mathbf{0}))$. From (i) and

$\forall w(w < x_{n+1} \rightarrow \neg A(\mathbf{p}, w, \mathbf{0}))$, we have $\neg \mathbf{i} < x_{n+1}$.

From $A(\mathbf{p}, x_{n+1}, \mathbf{0})$ and $(i+1)$, we have $\neg x_{n+1} < \mathbf{i}$. Thus by 14.13 we have $x_{n+1} = \mathbf{i}$. So $\vdash_Q \forall x_{n+1}(B(\mathbf{p}, x_{n+1}) \rightarrow x_{n+1} = \mathbf{i})$.

Exercises

14.1 Verify the following assertion: all recursive functions are representable in the theory ('R') whose language is L and whose theorems are the consequences in L of the following infinitely many sentences:

$\mathbf{i} \neq \mathbf{j}$ for all i, j such that $i \neq j$;

$\mathbf{i}+\mathbf{j} = \mathbf{k}$ for all i, j, k such that $i+j = k$;

$\mathbf{i} \cdot \mathbf{j} = \mathbf{k}$ for all i, j, k such that $i \cdot j = k$;

$\forall x(x < \mathbf{i} \rightarrow x = \mathbf{0} \vee \ldots \vee x = \mathbf{i}-\mathbf{1})$ for all i;

and $\forall x(x < \mathbf{i} \vee x = \mathbf{i} \vee \mathbf{i} < x)$, for all i.

14.2 Show that none of the following sentences are theorems of Q:

(*a*) $\forall x\, x \neq x'$,

(*b*) $\forall x \forall y \forall z\, x+(y+z) = (x+y)+z$,

(*c*) $\forall x \forall y\, x+y = y+x$,

(*d*) $\forall x\, \mathbf{0}+x = x$,

(*e*) $\forall x\, x < x'$,

(*f*) $\forall x \forall y \neg(x < y \,\&\, y < x)$,

(*g*) $\forall x \forall y \forall z\, x \cdot (y \cdot z) = (x \cdot y) \cdot z$,

(*h*) $\forall x \forall y\, x \cdot y = y \cdot x$,

(*i*) $\forall x\, \mathbf{0} \cdot x = \mathbf{0}$,

(*j*) $\forall x \forall y \forall z\, x \cdot (y+z) = x \cdot y + x \cdot z$.

Hint: Let a and b be two objects that are not natural numbers, and consider the following successor, addition, and multiplication tables:

x	x'
i	i'
a	a
b	b

$+$	j	a	b
i	$i+j$	b	a
a	a	b	a
b	b	b	a

$\cdot$	$\mathbf{0}$	$j \neq \mathbf{0}$	a	b
i	$\mathbf{0}$	$\mathbf{0}$	a	b
a	$i \neq \mathbf{0}$	$\mathbf{0}$	$i \cdot j$	a	b
b	a	$\mathbf{0}$	b	b	b
	b	$\mathbf{0}$	a	a	a

15
Undecidability, indefinability and incompleteness

We are now in a position to give a unified treatment of some of the central negative results of logic: Church's theorem on the undecidability of logic, Tarski's theorem on the indefinability of truth, and Gödel's first theorem on the incompleteness of systems of arithmetic. These theorems can all be seen as more or less direct consequences of the result of the last chapter, that all recursive functions are representable in Q, and a certain exceedingly ingenious lemma ('the diagonal lemma'), the idea of which is due to Gödel, and which we shall prove below. The first notion that we have to introduce is that of a *gödel numbering*.

A *gödel numbering* is an assignment of natural numbers (called 'gödel numbers') to expressions (in some set) that meets these conditions: (1) different gödel numbers are assigned to different expressions: (2) it is effectively calculable what the gödel number of any expression is; (3) it is effectively decidable whether a number is the gödel number of some expression in the set, and, if so, effectively calculable which expression it is the gödel number of.

Gödel numberings enable one to regard interpreted languages supposed to be 'about' the natural numbers – i.e. having the set of natural numbers as the domain of their intended interpretation – as also referring to the numbered expressions. The possibility then arises that certain sentences, ostensibly referring to certain numbers, could be seen as referring, via the gödel numbering, to certain expressions that are *identical* with those very sentences themselves. The state of affairs just described is no mere possibility; the proof of the diagonal lemma shows how it arises, and succeeding theorems show how it may be exploited.

We shall consider a particular set of expressions and a particular gödel numbering, to which we appropriate the words 'expression' and 'gödel number'. There is nothing special about our particular gödel numbering; the theorems and proofs that we are going to give with respect to the one we use could have been given with respect to any number of others. Our expressions are finite sequences of these (distinct) symbols.

We'll make the following 'conventions' about the identity of certain symbols: we stipulate that $x_0 = x$, $x_1 = y$, $f^0_0 = 0$, $f^1_0 = '$, $f^2_0 = +$, $f^2_1 = \cdot$,

TABLE 15-1

( )	&	∃	x_0	f^0_0	f^1_0	f^2_0	...	A^0_0	A^1_0	A^2_0	...
,	∨	∀	x_1	f^0_1	f^1_1	f^2_1	...	A^0_1	A^1_1	A^2_1	...
	–		x_2	f^0_2	f^1_2	f^2_2	...	A^0_2	A^1_2	A^2_2	...
	↔		.	.	.	.		.	.	.	
	→		.	.	.	.		.	.	.	
			.	.	.	.		.	.	.	

and $A^2_0 = \ =$. We now assign each symbol in Table 15-1 the number in the corresponding location in Table 15-2 as its gödel number:

TABLE 15-2

1	2	3	4	5	6	68	688	...	7	78	788	...
	29	39	49	59	69	689	6889	...	79	789	7889	...
		399		599	699	6899	68899	...	799	7899	78899	...
		3999		.	.	.	.		.	.	.	
		39999		.	.	.	.		.	.	.	
				.	.	.	.		.	.	.	

We'll write 'gn' to mean 'the gödel number of'. Thus,

$$\text{gn}(x) = 5,\ \text{gn}(y) = 59,\ \text{gn}(0) = 6,\ \text{gn}(') = 68,\ \text{gn}(+) = 688,$$
$$\text{gn}(\cdot) = 6889, \quad \text{and} \quad \text{gn}(=) = 788.$$

We must now extend the gödel numbering so that all finite sequences of symbols in Table 15-1 are assigned gödel numbers. (We don't distinguish between a single symbol and the sequence which consists of that one symbol.) The principle can be indicated in a single example: Since $\text{gn}(\exists) = 4$, $\text{gn}(x) = 5$, $\text{gn}(($ $) = 1$, and $\text{gn}(=) = 788$, we want

$$\text{gn}(\exists x(x =)$$

to be 4515788.

The principle is that if expression A has gödel number i, and B has j, then AB, the expression formed by writing A immediately before B, is to have as its gödel number the number denoted by the decimal arabic numeral formed by writing the decimal arabic numeral for i immediately before the decimal arabic numeral for j. It's clear that our gödel numbering really is a gödel numbering in the sense of the second paragraph.

We now introduce the notion of the *diagonalization* of an expression A. Recall from the last chapter that $\mathbf{n}$ is the result of writing n occurrences of $'$ immediately after $\mathbf{o}$. If A is an expression with gödel number n, we define $\ulcorner A \urcorner$ to be the expression $\mathbf{n}$. In what follows $\ulcorner A \urcorner$ will be seen to behave rather like a name for the expression A. The *diagonalization* of A is the expression

$$\exists x(x = \ulcorner A \urcorner \,\&\, A).$$

If A is a formula in the language of arithmetic that contains just the variable x free, then the diagonalization of A will be a sentence that 'says that' A is true of its own gödel number – or, more precisely, the diagonalization will be true (in the standard interpretation $\mathcal{N}$) if and only if A is true (in $\mathcal{N}$) of its own gödel number.

Lemma 1

There is a recursive function, diag, such that if n is the gödel number of an expression A, diag (n) is the gödel number of the diagonalization of A.

Proof. Let lh ('length') be defined by $\mathrm{lh}(n) = \mu m(\mathrm{o} < m \,\&\, n < \mathrm{10}^m)$. Since every natural number n is less than $\mathrm{10}^m$ for some positive m, and as exponentiation and less than are recursive, lh is a recursive function. lh (n) is the number of digits in the usual arabic numeral for n. Thus $\mathrm{lh}(1879) = 4$; $\mathrm{lh}(\mathrm{o}) = \mathrm{lh}(9) = 1$.

Let $*$ be defined by: $m * n = m \cdot \mathrm{10}^{\mathrm{lh}(n)} + n$. $*$ is recursive. If $m \neq \mathrm{o}$, $m*n$ is the number denoted by the arabic numeral formed by writing the arabic numeral for m immediately before the arabic numeral for n. Thus $2*3 = 23$.

Let num be defined by: $\mathrm{num}(\mathrm{o}) = 6$; $\mathrm{num}(n+1) = \mathrm{num}(n) * 68$ (all n). num is recursive. As $\mathrm{gn}(\mathbf{o}) = 6$ and $\mathrm{gn}(') = 68$, num (n) is the gödel number of $\mathbf{n}$.

As no arabic numeral for a gödel number contains the digit 'o', diag (n) may be taken to $= 4515788 * (\mathrm{num}(n) * (3 * (n * 2)))$. diag is then recursive.

We'll now reserve the word 'theory' for those theories whose variables are $x_0, x_1, \ldots$, and whose names, n-place function signs, sentence letters, and n-place predicate letters are some (or all) of $f_0^0, f_1^0, \ldots, f_0^n, f_1^n, \ldots$, $A_0^0, A_0^1, \ldots, A_0^n, A_1^n, \ldots$, respectively. As in the last chapter, we assume that $\mathbf{o}$ and $'$ occur in the language of all theories. We assume further that the set of gödel numbers of symbols of the language of the theory is recursive,

i.e. (by Church's thesis), that there is an effective procedure for deciding whether a given symbol may occur in some sentence in the language of the theory.

Here's the diagonal lemma:

Lemma 2

Let T be a theory in which diag is representable. Then for any formula $B(y)$ (of the language of T, containing just the variable y free), there is a sentence G such that

$$\vdash_T G \leftrightarrow B(\ulcorner G \urcorner).$$

Proof. Let $A(x, y)$ represent diag in T. Then for any n, k, if $\mathrm{diag}(n) = k$, $\vdash_T \forall y(A(\mathbf{n}, y) \leftrightarrow y = \mathbf{k})$.

Let F be the expression $\exists y(A(x,y) \,\&\, B(y))$. F is a formula of the language of T that contains just the variable x free.

Let n be the gödel number of F.

Let G be the expression $\exists x(x = \mathbf{n} \,\&\, \exists y(A(x,y) \,\&\, B(y)))$. As $\mathbf{n} = \ulcorner F \urcorner$, G is the diagonalization of F and a sentence of the language of T. Since G is logically equivalent to $\exists y(A(\mathbf{n}, y) \,\&\, B(y))$, we have

$$\vdash_T G \leftrightarrow \exists y(A(\mathbf{n}, y) \,\&\, B(y)).$$

Let k be the gödel number of G. Then

$$\mathrm{diag}(n) = k, \quad \text{and} \quad \mathbf{k} = \ulcorner G \urcorner.$$

So $\quad \vdash_T \forall y(A(\mathbf{n}, y) \leftrightarrow y = \mathbf{k})$.

So $\quad \vdash_T G \leftrightarrow \exists y(y = \mathbf{k} \,\&\, B(y))$.

So $\quad \vdash_T G \leftrightarrow B(\mathbf{k})$, i.e., $\vdash_T G \leftrightarrow B(\ulcorner G \urcorner)$.

A theory is called *consistent* if there is no theorem of the theory whose negation is also a theorem. Equivalently, a theory is consistent iff there is some sentence in its language that is not a theorem, iff the theory is satisfiable.

A set θ of natural numbers is said to be *definable in* theory T if there is a formula $B(x)$ of the language of T such that for any number k, if $k \in \theta$, then $\vdash_T B(\mathbf{k})$, and if $k \notin \theta$, then $\vdash_T -B(\mathbf{k})$. The formula $B(x)$ is said to define θ in T. A two-place relation R of natural numbers is likewise definable in T if there is a formula $C(x,y)$ of the language of T such that for any numbers k, n, if kRn, then $\vdash_T C(\mathbf{k}, \mathbf{n})$, and if $k\not{R}n$, then $\vdash_T -C(\mathbf{k}, \mathbf{n})$, and $C(x,y)$ is then said to define R in T. (A perfectly analogous definition

of definability can be given for three- and more-place relations on natural numbers; we won't need this more general notion, however.)

A theory T is called an *extension* of theory S if S is a subset of T, i.e., if any theorem of S is a theorem of T. If f is a function that is representable in S, and T is an extension of S, then f is representable in T, and indeed is represented in T by the same formula that represents it in S. Similarly, any formula that defines a set in some theory defines it in any extension of that theory.

Lemma 3

If T is a consistent extension of Q, then the set of gödel numbers of theorems of T is not definable in T.

Proof. Let T be an extension of Q. Then diag is representable in T; for as diag is a recursive function, and all recursive functions are representable in Q, diag is representable in Q, and hence is representable in any extension of Q.

Suppose now that $C(y)$ defines the set θ of gödel numbers of theorems of T. By the diagonal lemma, there is a sentence G such that

$$\vdash_T G \leftrightarrow -C(\ulcorner G \urcorner).$$

Let $k = \mathrm{gn}(G)$. Then

$$\vdash_T G \leftrightarrow -C(\mathbf{k}). \qquad (*)$$

Then $\vdash_T G$. For if G is not a theorem of T, then $k \notin \theta$, and so, as $C(y)$ defines θ, $\vdash_T -C(\mathbf{k})$, whence by (*), $\vdash_T G$.

So $k \in \theta$. So $\vdash_T C(\mathbf{k})$, as $C(y)$ defines θ. So, by (*), $\vdash_T -G$, and T is therefore inconsistent.

A set of expressions is called *decidable* if the set of gödel numbers of its members is a recursive set. Thus a theory T is decidable iff the set θ of gödel numbers of its theorems is recursive, iff the characteristic function of θ is recursive.

If a theory is decidable, then an effective method exists for deciding whether any given sentence is a theorem of the theory. For to determine whether a sentence is a theorem, calculate its gödel number first and then calculate the value of the (recursive, hence calculable) characteristic function for the gödel number as argument. The sentence is a theorem iff the value is 1.

Conversely, if a theory is not decidable, then *unless Church's thesis is false*, no effective method exists for deciding whether a given sentence is a theorem of the theory. For if there were such a method, then the characteristic function of the set of gödel numbers of theorems would also be effectively calculable, and hence recursive, by Church's thesis.

Theorem 1

No consistent extension of Q is decidable.

Proof. Suppose T is a consistent extension of Q. Then by Lemma 3, the set θ of gödel numbers of theorems of T is not definable in T. Now if $A(x,y)$ represented the characteristic function f of θ in T, then $A(x,1)$ would define θ in T. (For then if $k \in \theta$, $f(k) = 1$, whence $\vdash_T A(\mathbf{k},1)$; and if $k \notin \theta$, $f(k) = 0$, whence $\vdash_T \forall y(A(\mathbf{k},y) \leftrightarrow y = \mathbf{0})$, whence, as $\vdash_Q \mathbf{0} \neq 1$, $\vdash_T -A(\mathbf{k},1)$.) Thus the characteristic function of θ is not representable in T, and therefore, as T is an extension of Q, not representable in Q either, and hence not recursive. So T is not decidable.

Lemma 4

Q is not decidable.

Proof. Q is a consistent extension of Q.

We can now give another proof of the proposition that first-order logic has no decision procedure, a proof that is rather different from the one given in Chapter 10.

Let L be the theory in L, the language of arithmetic, whose theorems are just the valid sentences in L. All theorems of L are theorems of Q, of course, but as not all of (indeed, none of) the axioms of Q are valid, L is not an extension of Q, and we cannot therefore apply theorem 1. But because Q has only finitely many axioms, we can nonetheless prove that L is not decidable, and hence that there is no effective method for deciding whether or not a first-order sentence is valid.

Theorem 2 (*Church's undecidability theorem*)

L is not decidable.

Proof. Let C be a conjunction of the axioms of Q. Then a sentence A is a theorem of Q iff C implies A, iff $(C \rightarrow A)$ is valid, iff $(C \rightarrow A)$ is a

theorem of L. (So, intuitively, a test for validity would yield a test for theoremhood in Q: to decide whether A is a theorem of Q, test $(C \to A)$ for validity.)

Let q be the gödel number of C. Let f be defined by:

$$f(n) = 1*(q*(39999*(n*2))).$$

f is recursive. If n is the gödel number of A, then $f(n)$ is the gödel number of $(C \to A)$.

Let λ be the set of gödel numbers of theorems of L. If λ is recursive, then so is $\{n|f(n)\in\lambda\}$. But $\{n|f(n)\in\lambda\}$ is the set of gödel numbers of theorems of Q, which, by lemma 4, is not recursive. Thus λ is not recursive and L is not decidable.

By *arithmetic* we shall understand that theory whose language is L and whose theorems are just the sentences of L that are *true* in the standard interpretation $\mathcal{N}$, in which the domain is the set of all natural numbers, and **0**, $'$, $+$, and $\cdot$ are assigned zero, successor, addition, and multiplication, respectively.

Theorem 3

Arithmetic is not decidable.

Proof. Arithmetic is a consistent extension of Q, and by Theorem 1 no consistent extension of Q is decidable.

Thus, unless Church's thesis is false, there is no effective method for deciding whether an arbitrary sentence in the language of arithmetic is true or false in $\mathcal{N}$. This negative result is in contrast to Presburger's theorem, proved in Chapter 21, that an effective method exists for deciding whether an arbitrary sentence in the language of arithmetic *not containing* '$\cdot$' is true or false (in $\mathcal{N}$).

Theorem 4 (*Tarski's indefinability theorem*)

The set of gödel numbers of sentences true in $\mathcal{N}$ is not definable in arithmetic.

Proof. Since the theorems of arithmetic are just the sentences true in $\mathcal{N}$, Theorem 4 follows from Lemma 3.

As any formula $B(x)$ will be true (in $\mathcal{N}$) of the number k if and only if $B(\mathbf{k})$ is a theorem of arithmetic, another way to put Theorem 4 is to say

that there is no formula of the language of arithmetic (with one free variable) which is true of just those natural numbers that are gödel numbers of truths of arithmetic, or, more briefly, 'arithmetical truth is not arithmetically definable'.

Lemma 5

Any recursive set is definable in arithmetic.

Proof. Suppose θ is a recursive set. Then the characteristic function of θ is recursive, and hence representable in Q. As in the proof of Theorem 1, θ is then definable in Q, and hence definable in arithmetic, which is an extension of Q.

Lemma 5 shows that Theorem 4 is at least as strong a result as Theorem 3, as Theorem 3 says that the set of gödel numbers of truths of $\mathcal{N}$ is not recursive. Since the converse of Lemma 5 does not hold (cf. Exercise 3), Theorem 4 is actually stronger than Theorem 3.

A theory T is called *complete* if for every sentence A (in the language of T), either A or $-A$ is a theorem of T. A theory T is consistent and complete, then, iff for any sentence A, exactly one of A and $-A$ is a theorem. Arithmetic is a consistent, complete extension of Q.

A theory T is called *axiomatizable* if there is a decidable subset of T whose consequences (in the language of T) are just the theorems of T. If there is a finite, and hence decidable, subset with this property, the theory is said to be *finitely axiomatizable*. It is clear from the definition of axiomatizability that any decidable theory is axiomatizable; Q is an example of a (finitely) axiomatizable theory that is not decidable.

The version of Gödel's incompleteness theorem that we shall prove is the assertion that there is no complete, consistent, axiomatizable extension of Q. That there is none will follow from Theorem 1 and the proposition (Theorem 5) that any axiomatizable complete theory is decidable.

This last proposition should be confused neither with the statement that every complete decidable theory is axiomatizable, which is trivially true, nor with the statement that every decidable axiomatizable theory is complete, which is false (counterexample: the theory whose non-logical symbols are the sentence letters p and q, and whose theorems are the consequences in this language of p).

Theorem 5

Any axiomatizable complete theory is decidable.

Proof. Let T be any theory whatsoever. Since the set of symbols that may occur in sentences of the language of T is decidable (as we have assumed earlier), the set of sentences of the language of T is itself decidable, i.e. there exists a Turing machine which, when given a number as input, yields 1 as output iff the number is the gödel number of a sentence of the language of T, and yields 0 otherwise.

Now suppose that T is axiomatizable and complete. If T is inconsistent, then the theorems of T are just the sentences in the language of T, which, we have just noted, form a decidable set. We may therefore suppose that T is consistent also.

Since T is axiomatizable, there is a decidable set S of sentences whose consequences (in the language of T) are just the theorems of T. Let A be a sentence (in the language of T). We shall say that a sentence is *A-interesting* if it is a conditional of which the antecedent is a conjunction of members of S and the consequent is either A or $-A$. Then, A is a theorem of T iff there is a *valid* A-interesting sentence whose consequent is A itself. And, since T is consistent and complete, A is a theorem of T iff $-A$ is not a theorem of T, iff there is *no* valid A-interesting sentence whose consequent is $-A$. Since S is decidable, so is the set of A-interesting sentences (for each sentence A).

We shall show that T is decidable by showing that there exists a Turing machine M which, when given a sentence A of the language of T as input, yields 1 as output iff A is a theorem of T, and yields 0 as output otherwise.

In Chapter 12 we established the existence of a Turing machine M^* which, when given any sentence as input, terminated after a finite number of steps with the production of the words 'yes, valid' iff the given sentence was valid.

Our machine M works by first writing down the number 1 after the input sentence A and then going into a loop consisting of a sequence of *subroutines*. In the nth of these, M writes down those $k (\leqslant n)$ sentences with gödel numbers $\leqslant n$ that are A-interesting and then 'imitates' M^* k times, each time performing n steps in the operation of M^* when given as input (a new) one of the k A-interesting sentences that have been written down. If one or more of these k sentences is shown valid (by the production of the words 'yes, valid') after n such steps, M picks one of them and determines whether its consequent is A or $-A$ (as T is consistent, the case cannot arise in which one sentence has consequent A and another has $-A$), and then yields as output 1 or 0, accordingly.

But if not, M erases everything except A and the number after it, to which it adds 1, and then goes into the $n+1$st subroutine.

Since either A or $-A$ is a theorem of T, but not both, there is a valid A-interesting sentence C, with gödel number i, and M^*, when applied to C, will terminate with 'yes, valid' after some finite number of steps, say j. M, therefore, when applied to A, will go into at most $\max(i,j)$ subroutines before yielding 1 or 0 as output, and will yield 1 iff the consequent of C is A. So M yields 1 as output iff A is valid, and yields 0 otherwise. T is therefore decidable.

Theorem 6 (*Gödel's first incompleteness theorem*)

There is no consistent, complete, axiomatizable extension of Q.

Proof. Theorem 6 is an immediate consequence of Theorems 1 and 5.

Corollary

Arithmetic is not axiomatizable.

The import of Gödel's first incompleteness theorem is sometimes expressed in the words 'any sufficiently strong formal theory (or system) of arithmetic is incomplete (if it is consistent)'. A 'formal' theory may be taken to be one whose theorems are deducible via the usual rules of logic from an axiom system. Since an axiom system is here understood to be a set of sentences for which an effective procedure for determining membership exists, and since the usual rules of logic are sound and complete, that is, since all and only the logical consequences of a set of sentences can be deduced from the set by means of the rules, 'formal theory' can be considered synonymous with 'axiomatizable theory'. 'A formal theory of arithmetic' can therefore be taken to be an axiomatizable theory all of whose theorems are truths in some interpretation whose domain is the set of natural numbers and in which those of $\mathbf{0}$, $'$, $+$, $\cdot$, $<$, ^{2}, etc. that occur in the theorems have their familiar, standard meanings.

Theorem 6 thus represents a sharpening of the above statement of Gödel's theorem in that it indicates a sufficient condition for 'sufficient strength', *viz.*, *being an extension of* Q. Q, as we have seen, is a rather weak theory (cf. Exercise 14.2), and Theorem 6 is thus a correspondingly strong result. It follows from Theorem 6 that any consistent mathematical theory of which the theorems are just the consequences of some effectively specified set of axioms, and among which are the seven axioms of Q,

is incomplete; hence for any interpretation of the language of the theory there will be truths in that interpretation which are not theorems of the theory. And perhaps the most significant consequence of Theorem 6 is what it says about the notions of *truth* (in the standard interpretation for the language of arithmetic) and *theoremhood*, or *provability* (in any particular formal theory): *that they are in no sense the same.*

Exercises

15.1 A formula $B(y)$ is called a truth-predicate for T if for any sentence G of the language of T, $\vdash_T G \leftrightarrow B(\ulcorner G \urcorner)$. Show that if T is a consistent theory in which diag is representable, then there is no truth-predicate for T.

15.2 Show that all functions representable in Q are recursive.

15.3 A set S of natural numbers is called *recursively enumerable* (*r.e.*) if there is a (two-place) recursive relation R such that $S = \{x \mid \exists y\, Rxy\}$. Show that for any set S, S is recursive iff both S and $\bar{S}$ are r.e. (Kleene's theorem). Are all r.e. sets definable in arithmetic? (Yes. Why?) Give some examples of r.e. sets and some examples of non-r.e. sets.

15.4 (*Craig*) Show that a theory T is axiomatizable if T is r.e., i.e. if the set of gödel numbers of members of T is r.e.

15.5 Let $B_1(y)$ and $B_2(y)$ be two formulas of the language of T with y as sole free variable. Show how to construct sentences A_1 and A_2 such that $\vdash_T A_1 \leftrightarrow B_1(\ulcorner A_2 \urcorner)$ and $\vdash_T A_2 \leftrightarrow B_2(\ulcorner A_1 \urcorner)$.

Solution to 15.2 (*Using Church's thesis*)

15.2 Suppose $A(\mathsf{x}, y)$ represents f in Q. Since Q is consistent and $\mathbf{m} \neq \mathbf{n}$ is a theorem of Q whenever $m \neq n$, $\vdash_Q \forall y\,(A(\mathsf{p}, y) \leftrightarrow y = \mathbf{m})$ iff $f(\mathsf{p}) = m$. In order to calculate $f(\mathsf{p})$, then, one may use a 'search procedure' similar to the one used in the proof of Theorem 5 to determine for which m the conditional whose antecedent is some fixed conjunction of the axioms of Q and whose consequent is $\forall y\,(A(\mathsf{p}, y) \leftrightarrow y = \mathbf{m})$ is valid. That m – it will be unique – is $f(\mathsf{p})$.

Solution to 15.4 (very tricky)

Suppose that R is a recursive relation and

$$\{x \mid x \text{ is the gödel number of a member of } T\} = \{x \mid \exists y\, Rxy\}.$$

For any sentence A and natural number y, let A^y be the conjunction $(A\&\ldots(A\&A)\ldots)$ of $y+2$ occurrences of A. Thus, e.g.

$$A^2 = (A\&(A\&(A\&A)))$$

and $A^0 = (A\&A)$. Let $U = \{A^y \mid Rgn(A)y\}$. If $A \in T$, then for some y, $Rgn(A)y$ and $A^y \in U$; and if $A^y \in U$, then $A \in T$. Since A and A^y are equivalent, T and U imply the same sentences, and the set of sentences in the language of T that follow from U is thus T itself. To show that T is axiomatizable, then, we need only show that U is decidable. But U *is* decidable: to decide whether an arbitrary sentence B is in U, we may apply the following effective procedure. Determine whether B is the conjunction $(A\&\ldots(A\&A)\ldots)$ of z occurrences of some sentence A, for some $z \geqslant 2$. If not, $B \notin U$. But if so, find A and z, and let $x = \mathrm{gn}(A)$ and $y = z-2$. Determine whether Rxy. (R is recursive.) If so, $B \in U$; if not, $B \notin U$.