If T is a consistent theory in the language of arithmetic, we say a set S is *defined* in T by $D(x)$ if for all n, if n is in S, then $D(\mathbf{n})$ is a theorem of T, and if n is not in S, then $\sim D(\mathbf{n})$ is a theorem of T. S is *definable* in T if S is defined by some formula. Arithmetical definability is simply the special case where T is *true arithmetic*, the set of all correct sentences. The general notion of definability in a theory extends to relations, but definability of a function turns out to be less useful than a related notion. For the remainder of this chapter, unless otherwise noted, 'function' will mean 'total function'. Let f be a one-place function. (The definition we are about to give extends easily to many-place functions.) We say f is *representable* in T if there is a formula $F(x, y)$ such that whenever $f(a) = b$, the following is a theorem of T:

$$\forall y(F(\mathbf{a}, y) \leftrightarrow y = \mathbf{b}).$$

This is logically equivalent to the conjunction of the positive assertion

$$F(\mathbf{a}, \mathbf{b})$$

and the general negative assertion

$$\forall y(y \neq \mathbf{b} \rightarrow \sim F(\mathbf{a}, y)).$$

By contrast, definability would only require that we have the positive assertion and for each particular $c \neq b$ the relevant particular instance of the general negative assertion, namely, $\sim F(\mathbf{a}, \mathbf{c})$.

Now in the special case where T is true arithmetic, of course if each particular numerical instance is correct, then the universal generalization is correct as well, so representability and definability come to the same thing. But for other theories, each particular numerical instance may be a theorem without the universal generalization being a theorem, and representability is in general a stronger requirement than definability. Note that if T is a *weaker* theory than T^* (that is, if the set of theorems of T is a subset of the set of theorems of T^*), then the requirement that a function be representable in T is a *stronger* requirement than that it be representable in T^* (that is, representability in T implies representability in T^*). Thus far we have proved all recursive functions to be representable in true arithmetic. If we are to strengthen our results, we must consider weaker theories than that.

16.2 Minimal Arithmetic and Representability

We now introduce a finite set of *axioms of minimal arithmetic* **Q**, which, though not strong enough to prove major theorems of number theory, at least are correct and strong enough to prove all correct $\exists$-rudimentary sentences. By themselves, the axioms of **Q** would not be adequate for number theory, but any set of adequate axioms would have to include them, or at least to prove them (in which case the set might as well include them). Our main theorems (Theorems 16.13 and 16.15) apply to any theory T that contains **Q**, and since **Q** is weak, the theorems are correspondingly strong.

In displaying the list of axioms we make use of a traditional convention, according to which when displaying sentences of the language of arithmetic that begin with a

string of one or more universal quantifiers, one may omit to write the quantifiers and write only the open formula that comes after them.

(Q1) $\mathbf{0} \neq x'$

(Q2) $x' = y' \to x = y$

(Q3) $x + \mathbf{0} = x$

(Q4) $x + y' = (x + y)'$

(Q5) $x \cdot \mathbf{0} = \mathbf{0}$

(Q6) $x \cdot y' = (x \cdot y) + x$

(Q7) $\sim x < \mathbf{0}$

(Q8) $x < y' \leftrightarrow (x < y \vee x = y)$

(Q9) $\mathbf{0} < y \leftrightarrow y \neq \mathbf{0}$

(Q10) $x' < y \leftrightarrow (x < y \,\&\, y \neq x')$

Thus axiom (Q1) is really $\forall x\ \mathbf{0} \neq x'$, axiom (Q2) is really $\forall x \forall y\ (x' = y' \to x = y)$, and so on. As is said, the real axioms are the *universal closures* of the formulas displayed. The theory **Q** of *minimal arithmetic* is the set of all sentences of the language of arithmetic that are provable from (or, equivalently, are true in all models of) these axioms. The significance of the various axioms will become clear as we work through the steps of the proof of the main theorem of this section.

16.13 Theorem. An ∃-rudimentary sentence is correct if and only if it is a theorem of **Q**.

Proof: Since every axiom of **Q** is correct, so is every theorem of **Q**, and hence any ∃-rudimentary sentence provable from the axioms of **Q** is correct. All the work will go into proving the converse. To begin with zero and sucessor, for any natural number m, of course $\mathbf{m} = \mathbf{m}$ (where **m** is as always the numeral for m, that is, is the term $\mathbf{0}'^{\cdots'}$ with m accents $'$) is provable even without any axioms, by pure logic.

All of $\mathbf{0} \neq \mathbf{1}, \mathbf{0} \neq \mathbf{2}, \mathbf{0} \neq \mathbf{3}, \ldots$, are provable by (Q1) (since the numerals **1**, **2**, **3**, ... all end in accents). Then $\mathbf{1} = \mathbf{2} \to \mathbf{0} = \mathbf{1}, \mathbf{1} = \mathbf{3} \to \mathbf{0} = \mathbf{2}, \ldots$ are provable using (Q2), and since $\mathbf{0} \neq \mathbf{1}, \mathbf{0} \neq \mathbf{2}, \ldots$ are provable, it follows by pure logic that $\mathbf{1} \neq \mathbf{2}$, $\mathbf{1} \neq \mathbf{3}, \ldots$, are provable. Then $\mathbf{2} = \mathbf{3} \to \mathbf{1} = \mathbf{2}, \mathbf{2} = \mathbf{4} \to \mathbf{1} = \mathbf{3}, \ldots$ are provable, again using (Q2), and since $\mathbf{1} \neq \mathbf{2}, \mathbf{1} \neq \mathbf{3}, \ldots$, are provable, it follows by pure logic that $\mathbf{2} \neq \mathbf{3}, \mathbf{2} \neq \mathbf{4}, \ldots$ are provable. Continuing in the same way, if $m < n$, then $\mathbf{m} \neq \mathbf{n}$ is provable.

It follows by pure logic (the symmetry of identity) that if $m < n$, then $\mathbf{n} \neq \mathbf{m}$ is provable also. Since in general if $m \neq n$ we have either $m < n$ or $n < m$, it follows that if $m \neq n$ then both $\mathbf{m} \neq \mathbf{n}$ and $\mathbf{n} \neq \mathbf{m}$ are provable.

Turning now to order, note that using (Q8), $x < \mathbf{1} \leftrightarrow (x < \mathbf{0} \vee x = \mathbf{0})$ is provable, and (Q7) is $\sim x < \mathbf{0}$. By pure logic $x < \mathbf{1} \leftrightarrow x = \mathbf{0}$ is provable from these, so that $\mathbf{0} < \mathbf{1}$ is provable, and since we already know that $\mathbf{1} \neq \mathbf{0}, \mathbf{2} \neq \mathbf{0}, \ldots$ are provable, it follows that $\sim \mathbf{1} < \mathbf{1}, \sim \mathbf{2} < \mathbf{1}, \ldots$ are provable. Then using (Q8) again, $x < \mathbf{2} \leftrightarrow (x < \mathbf{1} \vee x = \mathbf{1})$ is provable, from which, given what we already know to be provable,

it follows that $x < \mathbf{2} \leftrightarrow (x = \mathbf{0} \vee x = \mathbf{1})$ is provable, from which it follows that $\mathbf{0} < \mathbf{2}$, $\mathbf{1} < \mathbf{2}$, and also $\sim\mathbf{2} < \mathbf{2}$, $\sim\mathbf{3} < \mathbf{2}$, ... are all provable. Continuing in the same way, for any m the following is provable:

$$x < \mathbf{m} \leftrightarrow (x = \mathbf{0} \vee x = \mathbf{1} \vee \ldots \vee x = \mathbf{m} - \mathbf{1}). \tag{1}$$

Moreover, whenever $n < m$, $\mathbf{n} < \mathbf{m}$ is provable, and whenever $m \geq n$, $\sim\mathbf{m} < \mathbf{n}$ is provable.

Turning now to addition and multiplication, let us show how (Q3) and (Q4), which are of course just the formal versions of the recursion equations for addition, can be used to prove, for instance, $\mathbf{2} + \mathbf{3} = \mathbf{5}$, or $\mathbf{0''} + \mathbf{0'''} = \mathbf{0'''''}$. Using (Q4), the following are all provable:

$$\mathbf{0''} + \mathbf{0'''} = (\mathbf{0''} + \mathbf{0''})'$$
$$\mathbf{0''} + \mathbf{0''} = (\mathbf{0''} + \mathbf{0'})'$$
$$\mathbf{0''} + \mathbf{0'} = (\mathbf{0''} + \mathbf{0})'.$$

Using (Q3), $\mathbf{0''} + \mathbf{0} = \mathbf{0''}$ is provable. Working backwards, by pure logic the following are all provable from what we have so far:

$$\mathbf{0''} + \mathbf{0'} = \mathbf{0'''}$$
$$\mathbf{0''} + \mathbf{0''} = \mathbf{0''''}$$
$$\mathbf{0''} + \mathbf{0'''} = \mathbf{0'''''}$$

This is, in fact, just the formal calculation exhibited in section 6.1. Obviously this method is perfectly general, and whenever $a + b = c$ we can prove $\mathbf{a} + \mathbf{b} = \mathbf{c}$. Then also, again as in section 6.1, the recursion equations (Q5) and (Q6) for multiplication can be used to prove $\mathbf{2} \cdot \mathbf{3} = \mathbf{6}$ and more generally, whenever $a \cdot b = c$ to prove $\mathbf{a} \cdot \mathbf{b} = \mathbf{c}$.

If we next consider more complex terms involving $'$ and $+$ and $\cdot$, their correct values are also provable. For example, consider $(\mathbf{1} + \mathbf{2}) \cdot (\mathbf{3} + \mathbf{4})$. By what we have already said, $\mathbf{1} + \mathbf{2} = \mathbf{3}$ and $\mathbf{3} + \mathbf{4} = \mathbf{7}$, as well as $\mathbf{3} \cdot \mathbf{7} = \mathbf{21}$, are provable. From these it is provable by pure logic that $(\mathbf{1} + \mathbf{2}) \cdot (\mathbf{3} + \mathbf{4}) = \mathbf{21}$, and similarly for other complex terms. Thus for any closed term t built up from $\mathbf{0}$ using $'$, $+$, $\cdot$, it is provable what is the correct value of the term. Suppose then we have two terms s, t that have the same value m. Since by what we have just said $s = \mathbf{m}$ and $t = \mathbf{m}$ are provable, by pure logic $s = t$ is also provable. Suppose instead the two terms have different values m and n. Then since $s = \mathbf{m}$ and $t = \mathbf{n}$ and $\mathbf{m} \neq \mathbf{n}$ are provable, again by pure logic $s \neq t$ is also provable. A similar argument applies to order, so all correct formulas of types $s = t$, $s \neq t$, $s < t$, $\sim s < t$ are provable. Thus all correct closed atomic and negated atomic sentences are provable.

Now we move beyond atomic and negation-atomic sentences. First, by pure logic the double negation of a sentence is provable if and only if the sentence itself is, and a conjunction is provable if both its conjuncts are, a disjunction is provable if either of its disjuncts is, a negated conjunction is provable if the negation of one of its conjuncts is, and a negated disjunction is provable if the negations of both of its disjuncts are. Since all correct atomic and negated atomic closed sentences are provable, so are all correct sentences of types $\sim S$, $\sim\sim S$, $S_1 \ \& \ S_2$, $\sim(S_1 \ \& \ S_2)$, $S_1 \vee S_2$, $\sim(S_1 \vee S_2)$,

where S, S_1, S_2 are atomic or negated atomic sentences. Continuing in this way, all correct closed formulas built up from atomic formulas by negation, conjunction, and disjunction are provable: All correct closed formulas without quantifiers are provable.

As for bounded quantifiers, using (1), for any formula $A(x)$ and any m, the following are provable:

$$\forall x < \mathbf{m} A(x) \leftrightarrow (A(\mathbf{0}) \,\&\, A(\mathbf{1}) \,\&\, \ldots \,\&\, A(\mathbf{m}-\mathbf{1})),$$
$$\exists x < \mathbf{m} A(x) \leftrightarrow (A(\mathbf{0}) \vee A(\mathbf{1}) \vee \ldots \vee A(\mathbf{m}-\mathbf{1})).$$

More generally, if t is a closed term whose correct value is m, since $t = \mathbf{m}$ is provable, so are the following:

$$\forall x < t A(x) \leftrightarrow (A(\mathbf{0}) \,\&\, A(\mathbf{1}) \,\&\, \ldots \,\&\, A(\mathbf{m}-\mathbf{1})),$$
$$\exists x < t A(x) \leftrightarrow (A(\mathbf{0}) \vee A(\mathbf{1}) \vee \ldots \vee A(\mathbf{m}-\mathbf{1})).$$

Thus any bounded universal or existential quantification of formulas without quantifiers can be proved equivalent to a conjunction or disjunction of sentences without quantifiers, which is of course itself then a sentence without quantifiers, so that we already know it can be proved if it is correct. Thus any correct sentence obtained by applying bounded universal or existential quantification to formulas without quantifiers is provable, and repeating the argument, so is any correct sentence built up from atomic formulas using negation, conjunction, disjunction, and bounded universal and bounded existential quantification: Any correct rudimentary sentence is provable.

Finally, consider now a correct $\exists$-rudimentary sentence $\exists x\, A(x)$. Since it is correct, there is some a such that $A(\mathbf{a})$ is correct. Being correct and rudimentary, $A(\mathbf{a})$ is provable, and hence so is $\exists x\, A(x)$, completing the proof.

Note that for a correct $\forall$-rudimentary sentence $\forall x\, A(x)$, we can conclude that each numerical instance $A(\mathbf{0})$, $A(\mathbf{1})$, $A(\mathbf{2})$, ... is provable from the axioms of **Q**, but this is not to say that $\forall x A(x)$ itself is provable from the axioms of **Q**, and in general it is not. There are nonstandard interpretations of the language of arithmetic on which all the axioms of **Q** come out true, but some very simple $\forall$-universal sentences that are correct or true on the standard interpretation come out false. Works on set theory develop an extremely natural nonstandard model of **Q**, called the system of *ordinal numbers,* for which, among others, laws as simple as $\mathbf{1} + x = x + \mathbf{1}$ fail. It would take us too far afield to stop to develop this model here, but some of its features are hinted at by the nonstandard interpretations of **Q** indicated in the problems at the end of the chapter. As we have already said, the fact that **Q** is a weak theory makes the following theorem (which automatically applies to any theory T containing **Q**) a strong theorem.

16.14 Lemma. Every rudimentary function is representable in **Q** (and by a rudimentary formula).

Proof: Inspection of the proof of the preceding lemma shows that it actually did not require any use of (Q9) and (Q10), but the proof of the present lemma does. An argument exactly like that used in the earlier proof to derive

(1) $$x < \mathbf{m} \leftrightarrow (x = \mathbf{0} \vee x = \mathbf{1} \vee \ldots \vee x = \mathbf{m}-\mathbf{1})$$

from (Q7) and (Q8) can be used to derive

$$\mathbf{m} < y \leftrightarrow (y \neq \mathbf{0} \,\&\, y \neq \mathbf{1} \,\&\, \ldots \,\&\, y \neq \mathbf{m}) \tag{2}$$

from (Q9) and (Q10). An immediate consequence of (1) and (2) together is the following:

$$z < \mathbf{m} \vee z = \mathbf{m} \vee \mathbf{m} < z. \tag{3}$$

Now let f be a one-place rudimentary function. (The proof for many-place functions is exactly the same.) Let $\phi(x, y)$ be a rudimentary formula arithmetically defining f. We do *not* claim that ϕ represents f in **Q**, but we do claim that ϕ can be used to build *another* rudimentary formula ψ that *does* represent f in **Q**. The formula $\psi(x, y)$ is simply

$$\phi(x, y) \,\&\, \forall z < y \sim \phi(x, z).$$

To show this formula represents f we must do two things. First, we must show that if $f(a) = b$, then $\psi(\mathbf{a}, \mathbf{b})$ is a theorem of **Q**. But indeed, since ϕ arithmetically defines f, if $f(a) = b$, then $\phi(\mathbf{a}, \mathbf{b})$ is correct, and $\sim\phi(\mathbf{a}, \mathbf{c})$ is correct for every $c \neq b$, and in particular for every $c < b$. Therefore $\forall z < \mathbf{b} \sim\phi(\mathbf{a}, z)$ is correct and $\psi(\mathbf{a}, \mathbf{b})$ is correct, and being rudimentary, it is a theorem of **Q** by Theorem 16.13.

Second, we must show that the following is a theorem of **Q**:

$$y \neq \mathbf{b} \rightarrow \sim\psi(\mathbf{a}, y),$$

which is to say

$$y \neq \mathbf{b} \rightarrow \sim(\phi(\mathbf{a}, y) \,\&\, \forall z < y \sim \phi(\mathbf{a}, z))$$

or, what is logically equivalent,

$$\phi(\mathbf{a}, y) \rightarrow (y = \mathbf{b} \vee \exists z < y \, \phi(\mathbf{a}, z)). \tag{4}$$

It will be sufficent to show that the following is a theorem of **Q**, since together with $\phi(\mathbf{a}, \mathbf{b})$, which we know to be a theorem of **Q**, it logically implies (4):

$$\phi(\mathbf{a}, y) \rightarrow (y = \mathbf{b} \vee \mathbf{b} < y). \tag{5}$$

But (3), together with $\forall y < \mathbf{b} \sim \phi(\mathbf{a}, y)$, which we know to be a theorem of **Q**, logically implies (5), to compete the proof.

16.15 Lemma. Any composition of rudimentary functions is representable in **Q** (and by an $\exists$-rudimentary formula).

Proof: We consider the composition of two one-place functions, the proof for many-place functions being similar. Suppose f and g are rudimentary functions, represented in **Q** by the rudimentary formulas ϕ_f and ϕ_g respectively. Let $h(x) = g(f(x))$, and consider the ($\exists$-rudimentary) formula ϕ_h we get from the proof of Lemma 16.4:

$$\exists y(\phi_f(x, y) \,\&\, \phi_g(y, z)) .$$

We claim ϕ_h represents h in **Q**. For let a be any number, $b = f(a)$, and $c = h(a) = g(f(a)) = g(b)$. Since ϕ_f represents f and $f(a) = b$, the following is a theorem of **Q**:

$$\forall y(\phi_f(\mathbf{a}, y) \leftrightarrow y = \mathbf{b})\,. \tag{1}$$

Since ϕ_g represents g and $g(b) = c$, the following is a theorem of **Q**:

$$\forall z(\phi_g(\mathbf{b}, z) \leftrightarrow z = \mathbf{c})\,. \tag{2}$$

What we need to show in order to establish that ϕ_h represents h in **Q** is that the following is a theorem of **Q**:

$$\forall z(\exists y(\phi_f(\mathbf{a}, y) \,\&\, \phi_g(y, z)) \leftrightarrow z = \mathbf{c})\,. \tag{3}$$

But (3) is logically implied by (1) and (2)!

16.16 Theorem

(a) Every recursive function is representable in **Q** (and by an ∃-rudimentary formula).
(b) Every recursive relation is definable in **Q** (and by an ∃-rudimentary formula).

Proof: (*a*) is immediate from Lemmas 16.12, 16.14, and 16.15. For (*b*), we consider the case of a one-place relation or set, many-place relations being similar. Let P be the recursive set, f its characteristic function, and $\exists w\phi(x, y, w)$ an ∃-rudimentary formula representing f in **Q**. If n is in P, then $f(n) = 1$, and **Q** proves $\exists w\phi(\mathbf{n}, \mathbf{1}, w)$. If n is not in P, then $f(n) = 0$, and **Q** proves $\forall y(y \neq \mathbf{0} \rightarrow \sim\exists w\phi(\mathbf{n}, y, w))$ and in particular $\sim\exists w\phi(\mathbf{n}, \mathbf{0}, w)$. So the formula $\exists w\phi(x, \mathbf{1}, w)$ defines P in **Q**.

Careful review of the proof of Theorem 16.16(a) shows that it actually applies to any recursive total *or partial* function f and gives a formula that *both arithmetically defines and* represents f in **Q**. This refinement will not be needed, however, for our work in the next chapter.

We now have all the machinery we need for the proof of the *first Gödel incompleteness theorem*, and readers impatient to see that famous result may skip ahead to the next chapter. They should then return to the next brief section of this one before going on to the *second Gödel incompleteness theorem* in the chapter after next.

16.3 Mathematical Induction

The most immediate reason for the inadequacy of the axioms of minimal arithmetic to prove many correct ∀-universal sentences is that they make no provision for proof by mathematical induction, a method ubiquitously used in number theory and mathematics generally, according to which we can prove that every number has some property by proving that zero has it (the *zero* or *basis* step), and proving that, assuming a number x has it (an assumption called the *induction hypothesis*) then the successor of x also has it (the *successor* or *induction* step).

16.17 Example (Dichotomy). As the most trivial example, we can prove by mathematical induction that every x is either 0 or the successor of some number. *Basis*. 0 is 0. *Induction.* x' is the successor of x.

Another example is the proof of the law

$$0+1+2+\cdots+x=x(x+1)/2.$$

Basis. $0=0\cdot 1/2$. *Induction*. Assuming the result for x, we have

$$\begin{aligned}0+1+2+\cdots+x+(x+1) &= x(x+1)/2+(x+1)\\ &= [x(x+1)+2(x+1)]/2\\ &= (x+1)(x+2)/2.\end{aligned}$$

The algebraic manipulations in this proof depend on basic laws of arithmetic (associative, commutative, distributive) which can be proved using mathematical induction.

16.18 Example (Additive identity). By mathematical induction one can prove (from the recursion equations defining addition) $0+x=x+0$. *Zero* or *basis* step: for $x=0$ we have $0+0=0+0$ by pure logic. *Successor* or *induction* step: assuming $0+x=x+0$, we have

$0+x'=(0+x)'$	by the second recursion equation for addition
$(0+x)'=(x+0)'$	by our assumption
$(x+0)'=x'=x'+0$	by the first recursion equation for addition.

16.19 Example (First case of the commutativity of addition). Similarly, we can prove $1+x=x+1$, or $0'+x=x+0'$. *Basis*: $0'+0=0+0'$ by the preceding example. *Induction*: assuming $0'+x=x+0'$, we have

$0'+x'=(0'+x)'$	by the second recursion equation for addition
$(0'+x)'=(x+0')'$	by assumption
$(x+0')'=(x+0)''$	by the second recursion equation for addition
$(x+0)''=x''$	by the first recursion equation for addition
$x''=(x'+0)'$	by the first recursion equation for addition
$(x'+0)'=x'+0'$	by the second recursion equation for addition.

We relegate further examples of this kind to the problems at the end of the chapter.

Once we have the basic laws of arithmetic, we can go on to prove various elementary lemmas of number theory such as the facts that a divisor of a divisor of a number is a divisor of that number, that every number has a prime factor, that if a prime divides a product it divides one of its factors, and that if two numbers with no common prime factor both divide a number, then so does their product. (The reader may recognize these as results we took for granted in the proof of Lemma 16.5.) Once we have enough elementary lemmas, we can go on to prove more substantial theorems of number theory, such as Lagrange's theorem from Example 16.7.

Closely related to the principle of mathematical induction as stated above is the principle of *complete induction*, according to which we can prove that every number has some property P by proving that zero has P, and proving that, assuming every number $\leq x$ has P, then the successor of x also has P. Indeed, complete induction for a property P follows on applying mathematical induction to the related property 'every number $\leq x$ has P,' using the facts (Q7) that 0 is the only number ≤ 0, and (Q8) that the only numbers $\leq x'$ are the numbers $\leq x$ and x' itself.

Another related principle is the *least-number principle*, according to which, if there is some number that has a given property, then there is a *least* number having the property, one such that no lesser number has it. This principle follows from the principle of mathematical induction as follows. Consider some property P such that there is *no* least number with the property P. Then we can use induction to show that in fact no number has the property P. We do this a bit indirectly, showing first by induction that for any number x, there is no number less than x with the property P. *Basis*: there is no number less than zero with the property P, because by (Q7) there is no number less than zero at all. *Induction*: supposing there is no number less than x with the property P, there can be no number less than the successor of x with the property P, since by (Q8) the only numbers less than the successor of x are the numbers less than x, which by assumption do not have the property, and x itself, which if it had the property would be the *least* number having the property. Now that we know that for any number x there is no number y less than x with the property, it follows that there is no number y with the property, since, taking x to be the successor of y, y is less than x and therefore cannot have the property.

(Conversely, the least-number principle together with the dichotomy of Example 16.17 yields the principle of mathematical induction. For if zero has a property and the successor of any number having the property has it also, then neither zero nor any successor can be the *least* number failing to have the property.)

All our argumentation in this section so far has been informal. A more adequate set of formal axioms for number theory is provided by the set of *axioms of Peano arithmetic*—an infinite (but primitive recursive) set of axioms consisting of the axioms of **Q** plus all sentences of the following form:

$$(A(\mathbf{0}) \,\&\, \forall x(A(x) \to A(x'))) \to \forall x\, A(x).$$

[Here $A(x)$ may contain other free variables $y_1, \ldots, y_n$, and what is really meant is

$$\forall y_1 \ldots \forall y_n((A(\mathbf{0}, y_1, \ldots, y_n) \,\&\, \forall x(A(x, y_1, \ldots, y_n) \to A(x', y_1, \ldots, y_n))) \to \forall x\, A(x, y_1, \ldots, y_n))$$

in accordance with the traditional convention of suppressing initial universal quantifiers in displayed formulas.]

The theory **P** of *Peano arithmetic* is the set of all sentences of the language of arithmetic that are provable from (or equivalently, are consequences of) these axioms. A rule to the effect that all sentences of a certain kind are to be taken as axioms is called an *axiom scheme*. With this terminology it would be said that the axioms of Peano arithmetic **P** consist of finitely many individual axioms (those of minimal arithmetic **Q**) plus a single axiom scheme (the *induction scheme* as above). In practice, the sets of axioms of most interest to logicians tend to consist of at most a dozen or so individual axioms and at most a very few axiom schemes, and so in particular are primitive recursive.

Among the axioms of **P** are for instance the following:

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

and

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

And using these axioms in addition to the axioms of **Q**, the laws $\mathbf{0}+x=x+\mathbf{0}$ and $\mathbf{1}+x=x+\mathbf{1}$ are provable from the axioms of **P**, by 'formalizing' the proof of these laws given above as Examples 16.18 and 16.19. Also, for any formula $F(x)$ the least-number principle for F, namely

$$\exists x F(x)\rightarrow \exists x(F(x)\ \&\ \forall y<x\sim F(y))$$

is provable from the axioms of **P**, again by 'formalizing' the proof given above; and similarly for complete induction. Eventually the usual proofs of, say, Lagrange's theorem in textbooks on number theory can be 'formalized' to give proofs from the axioms of **P**.

The method of proof by mathematical induction is indeed an ingredient in the proofs of essentially all major theorems in mathematics, but it is perhaps especially common in *metamathematics*, the branch of mathematics concerned with giving proofs *about* what can be proved in mathematics—the branch to which the present book belongs. We have been using this method of proof all along, often in disguise. Consider, for instance, the proof by induction on complexity of formulas, of which we have made considerable use. What one does with this method is, not to put too fine a point on it, prove (as base step) that any atomic formula, which is to say, any formula containing 0 occurrences of the logical symbols (negation, junctions, quantifiers), has a certain property, and then prove (as inductive step) that if all formulas containing no more than n occurrences of the logical symbols have the property, then so does any formula containing n' such occurrences. The proof of the latter assertion is broken down into cases according as the one extra symbol is a negation, a junction, or a quantifier. This method of proof is really a special form of proof by mathematical induction.

And in our proof of Theorem 16.13 in the preceding section, for instance, every step involved some sort of induction, though we have expressed it very casually, using such phrases as 'continuing in the same way'. A less casual way of putting the second paragraph of the proof, for instance, would be as follows:

It can be proved by mathematical induction that if $m<n$, then $\mathbf{m}\neq\mathbf{n}$ is provable from the axioms of **Q**. *Basis*: if $0<n$ then $\mathbf{0}\neq\mathbf{n}$ is provable by (Q0) (since the numeral **n** ends in an accent). *Induction*: assuming $\mathbf{m}\neq\mathbf{n}$ is provable whenever $m<n$, if $m'<n$, then we show $\mathbf{m'}\neq\mathbf{n}$ is provable as follows. Let $n=k'$. Then $m<k$, and by assumption $\mathbf{m}\neq\mathbf{k}$ is provable. But $\mathbf{m'}=\mathbf{k'}\rightarrow\mathbf{m}=\mathbf{k}$, which is to say $\mathbf{m'}=\mathbf{n}\rightarrow\mathbf{m}=\mathbf{k}$, is provable by (Q1). It follows by pure logic that $\mathbf{m}\neq\mathbf{n}$ is provable.

In this example we are using induction ('in the metalanguage') to prove something about a theory that does not have induction as an axiom ('in the object language'): we prove that something is a theorem of **Q** for every m by proving it is a theorem for 0, and that if it is a theorem for m, then it is a theorem for m'. Again, this sort of proof can be 'formalized' in **P**.

16.4* Robinson Arithmetic

This optional section is addressed to readers who wish to compare our treatment of the matters with which we have been concerned in this chapter with other treatments in the literature. In the literature, the label **Q** is often used to refer not to our minimal arithmetic but to another system, called *Robinson arithmetic*, for which we use the label **R**. To obtain the axioms of **R** from those of **Q**, add

(Q0) $$x = \mathbf{0} \vee \exists y\, x = y'$$

and replace (Q7)–(Q10) by

(Q11) $$x < y \leftrightarrow \exists z(z' + x = y).$$

We have already mentioned an extremely natural nonstandard model for **Q**, called the system of ordinal numbers, in which (Q0) fails. There is also an extremely natural nonstandard model for **R**, called the system of *cardinal numbers*, in which (Q10) fails; though it would take us too far afield to develop this model here, a simplified version suffices to show that some theorems of **Q** are not theorems of **R**. Thus **Q** is in some respects weaker and in some respects stronger than **R**, and vice versa.

By Theorem 16.16, every recursive function is representable in **Q**. Careful rereading of the proof reveals that all the facts it required about order are these, that the following are theorems:

(1) $$\mathbf{a} < \mathbf{b}, \text{ whenever } a < b$$

(2) $$\sim x < \mathbf{0}$$

(3) $$\mathbf{0} < y \leftrightarrow y \neq \mathbf{0}$$

and for any b the following:

(4) $$x < \mathbf{b}' \rightarrow x < \mathbf{b} \vee x = \mathbf{b}$$

(5) $$\mathbf{b} < y \,\&\, y \neq \mathbf{b}' \rightarrow \mathbf{b}' < y.$$

Clearly (1) is a theorem of **R**, since if $a < b$, then for some c, $c' + a = b$, and then $\mathbf{c'} + \mathbf{a} = \mathbf{b}$ is a consequence of (Q1)–(Q4). Also (2), which is axiom (Q7), is a theorem of **R**. For first $z' + \mathbf{0} = z' \neq \mathbf{0}$ by (Q3) and (Q1), which gives us $\sim \mathbf{0} < \mathbf{0}$,

and then second $z' + y' = (z' + y') \neq \mathbf{0}$ by (Q4) and (Q1), which gives us $\sim y' < \mathbf{0}$. But these two, together with (Q0), give us (2). Also (3), which is axiom (Q9), is a theorem of $\mathbf{R}$. For $\mathbf{0} < y \rightarrow y \neq \mathbf{0}$ follows from $\sim \mathbf{0} < \mathbf{0}$, and for the opposite direction, (Q0) gives $y \neq \mathbf{0} \rightarrow \exists z(y = z')$, while (Q3) gives $y = z' \rightarrow z' + \mathbf{0} = y$, and (Q11) gives $\exists z(z' + \mathbf{0} = y) \rightarrow \mathbf{0} < y$, and (3) is a logical consequence of these three.

It turns out that (4) and (5) are also theorems of $\mathbf{R}$, and hence every recursive function is representable in $\mathbf{R}$. Proofs have been relegated to the problems at the end of the chapter because we do not need any results about $\mathbf{R}$ for our later work. All we need for the purposes of proving the celebrated Gödel incompleteness theorems and their attendant lemmas and corollaries in the next chapter is that there is *some* correct, finitely axiomatizable theory in the language of arithmetic in which all recursive functions are representable. We chose minimal arithmetic because it is easier to prove representability for it; except in this regard Robinson arithmetic would really have done no worse and no better.

Problems

16.1 Show that the class of arithmetical relations is closed under substitution of recursive total functions. In other words, if P is an arithmetical set and f a recursive total function, and if $Q(x) \leftrightarrow P(f(x))$, then Q is an arithmetical set, and similarly for n-place relations and functions.

16.2 Show that the class of arithmetical relations is closed under negation, conjunction, disjunction, and universal and existential quantification, and in particular that every semirecursive relation is arithmetical.

16.3 A theory T is inconsistent if for some sentence A, both A and $\sim A$ are theorems of T. A theory T in the language of arithmetic is called *ω-inconsistent* if for some formula $F(x)$, $\exists x F(x)$ is a theorem of T, but so is $\sim F(\mathbf{n})$ for each natural number n. Let T be a theory in the language of arithmetic extending $\mathbf{Q}$. Show:

(a) If T proves any incorrect $\forall$-rudimentary sentence, then T is inconsistent.

(b) If T proves any incorrect $\exists$-rudimentary sentence, then T is ω-inconsistent.

16.4 Extend Theorem 16.3 to generalized $\exists$-rudimentary sentences.

16.5 Let R be the set of triples (m, a, b) such that m codes a formula $\phi(x, y)$ and $\mathbf{Q}$ proves

$$\forall y(\phi(\mathbf{a}, y) \leftrightarrow y = \mathbf{b}).$$

Show that R is semirecursive.

16.6 For R as in the preceding problem, show that R is the graph of a two-place partial function.

16.7 A *universal function* is a two-place recursive partial function F such that for any one-place recursive total or partial function f there is an m such that $f(a) = F(m, a)$ for all a. Show that a universal function exists.

The result of the preceding problem was already proved in a completely different way (using the theory of Turing machines) in Chapter 8 *as Theorem* 8.5. *After completing the preceding problem, readers who skipped section* 8.3 *may turn to it, and to the related problems at the end of Chapter* 8.

16.8 A set P is (*positively*) *semidefinable* in a theory T by a formula $\phi(x)$ if for every n, $\phi(\mathbf{n})$ is a theorem of T if and only if n is in P. Show that every semirecursive set is (positively) semidefinable in $\mathbf{Q}$ and any ω-consistent extension of $\mathbf{Q}$.

16.9 Let T be a consistent, axiomatizable theory containing $\mathbf{Q}$. Show that:

(a) Every set (positively) semi-definable in T is semirecursive.

(b) Every set definable in T is recursive.

(c) Every total function representable in T is recursive.

16.10 Using the recursion equations for addition, prove:

(a) $x + (y + 0) = (x + y) + 0$.

(b) $x + (y + z) = (x + y) + z \rightarrow x + (y + z') = (x + y) + z'$.

The *associative law* for addition,

$$x + (y + z) = (x + y) + z$$

then follows by mathematical induction ('on z'). (You may argue informally, as at the beginning of section 16.3. The proofs can be 'formalized' in **P**, but we are not asking you to do so.)

16.11 Continuing the preceding problem, prove:

(c) $x' + y = (x + y)'$,

(d) $x + y = y + x$.

The latter is the *commutative law* for addition.

16.12 Continuing the preceding problems, prove the *associative* and *distributive* and *commutative laws* for multiplication:

(e) $x \cdot (y + z) = x \cdot y + x \cdot z$,

(f) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$,

(g) $x \cdot y = y \cdot x$.

16.13 **(a)** Consider the following nonstandard order relation on the natural numbers: $m <_1 n$ if and only if m is odd and n is even, or m and n have the same parity (are both odd or both even) and $m < n$. Show that if there is a natural number having a property P then there is a $<_1$-least such natural number.

(b) Consider the following order on pairs of natural numbers: $(a, b) <_2 (c, d)$ if and only if either $a < c$ or both $a = c$ and $b < d$. Show that if there is a pair of natural numbers having a property P then there is a $<_2$-least such pair.

(c) Consider the following order on finite sequences of natural numbers: $(a_0, \ldots, a_m) <_3 (b_0, \ldots, b_n)$ if and only if either $m < n$ or both $m = n$ and the following condition holds: that either $a_m < b_m$ or else for some $i < m$, $a_i < b_i$ while for $j > i$ we have $a_j = b_j$. Show that if there is a sequence of natural numbers having a property P then there is a $<_3$-least such sequence.

16.14 Consider a nonstandard interpretation of the language $\{\mathbf{0}, ', <\}$ in which the domain is the set of natural numbers, but the denotation of $<$ is taken to be the

relation $<_1$ of Problem 16.13(a). Show that by giving suitable denotations to $\mathbf{0}$ and $'$, axioms (Q1)–(Q2) and (Q7)–(Q10) of **Q** can be made true, while the sentence $\forall x(x = \mathbf{0} \vee \exists y\ x = y')$ is made false.

16.15 Consider a nonstandard interpretation of the language $\{\mathbf{0}, ',\ <, +\}$ in which the domain is the set of pairs of natural numbers, and the denotation of $<$ is taken to be the relation $<_2$ of Problem 16.13(b). Show that by giving suitable denotations to $\mathbf{0}$ and $'$ and $+$, axioms (Q1)–(Q4) and (Q7)–(Q10) of **Q** can be made true, while both the sentence of the preceding problem and the sentence $\forall y(\mathbf{1} + y = y + \mathbf{1})$ are made false.

16.16 Consider an interpretation of the language $\{\mathbf{0}, ', +, \cdot,\ <\}$ in which the domain is the set of natural numbers plus one additional object called ∞, where the relations and operations on natural numbers are as usual, $\infty' = \infty$, $x + \infty = \infty + x = \infty$ for any x, $0 \cdot \infty = \infty \cdot 0 = 0$ but $x \cdot \infty = \infty \cdot x = \infty$ for any $x \neq 0$, and $x < \infty$ for all x, but not $\infty < y$ for any $y \neq \infty$. Show that axioms (Q0)–(Q9) and (Q11) are true on this interpretation, but not axiom (Q10).

16.17 Show that, as asserted in the proof of Lemma 16.14, for each m the following is a theorem of **Q**:

$$\mathbf{m} < y \leftrightarrow (y \neq \mathbf{0} \,\&\, y \neq \mathbf{1} \,\&\, \ldots \,\&\, y \neq \mathbf{m}).$$

16.18 Show that if the induction axioms are added to (Q1)–(Q8), then (Q9) and (Q10) become theorems. *The following problems pertain to the optional section 16.4.*

16.19 Show that the following are theorems of **R** for any b:

(a) $x' + \mathbf{b} = x + \mathbf{b}'$.

(b) $\mathbf{b} < x \rightarrow \mathbf{b}' < x'$.

(c) $x' < y' \rightarrow x < y$.

16.20 Show that the following are theorems of **R** for any b:

(a) $x < \mathbf{b}' \rightarrow x < \mathbf{b} \vee x = \mathbf{b}$.

(b) $\mathbf{b} < y \,\&\, y \neq \mathbf{b}' \rightarrow \mathbf{b}' < y$.

16.21 Show that adding induction to **R** produces the same theory (Peano arithmetic **P**) as adding induction to **Q**.