Graduate Texts in Mathematics

Donald W. Barnes
John M. Mack

An Algebraic Introduction to Mathematical Logic

Springer-Verlag
New York Heidelberg Berlin

Graduate Texts in Mathematics 22

Donald W. Barnes John M. Mack

An Algebraic Introduction to Mathematical Logic

Springer-Verlag New York Heidelberg Berlin

Donald W. Barnes
John M. Mack [redacted]
The University of Sydney
Department of Pure Mathematics
Sydney, N.S.W. 2006
Australia

AMS Subject Classifications
Primary: 0201
Secondary: 02B05, 02B10, 02F15, 02G05, 02G10, 02G15, 02G20, 02H05, 02H13, 02H15, 02H20, 02H25

Library of Congress Cataloging in Publication Data

Barnes, Donald W
An algebraic introduction to mathematical logic.
(Graduate texts in mathematics; v. 22)
Bibliography: p. 115
Includes index.
1. Logic, Symbolic and mathematical. 2. Algebraic
logic. I. Mack, J. M., joint author. II. Title.
III. Series.
QA9.B27 511'.3 74-22241

Printed in the United States of America.

ISBN 0-387-90109-4 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90109-4 Springer-Verlag Berlin Heidelberg New York

Preface

This book is intended for mathematicians. Its origins lie in a course of lectures given by an algebraist to a class which had just completed a substantial course on abstract algebra. Consequently, our treatment of the subject is algebraic. Although we assume a reasonable level of sophistication in algebra, the text requires little more than the basic notions of group, ring, module, etc. A more detailed knowledge of algebra is required for some of the exercises. We also assume a familiarity with the main ideas of set theory, including cardinal numbers and Zorn's Lemma.

In this book, we carry out a mathematical study of the logic used in mathematics. We do this by constructing a mathematical model of logic and applying mathematics to analyse the properties of the model. We therefore regard all our existing knowledge of mathematics as being applicable to the analysis of the model, and in particular we accept set theory as part of the meta-language. We are not attempting to construct a foundation on which all mathematics is to be based—rather, any conclusions to be drawn about the foundations of mathematics come only by analogy with the model, and are to be regarded in much the same way as the conclusions drawn from any scientific theory.

The construction of our model is greatly simplified by our using universal algebra in a way which enables us to dispense with the usual discussion of essentially notational questions about well-formed formulae. All questions and constructions relating to the set of well-formed formulae are handled by our Theorems 2.2 and 4.3 of Chapter I. Our use of universal algebra also provides us with a convenient method for discussing free variables (and avoiding reference to bound variables), and it also permits a simple neat statement of the Substitution Theorem (Theorems 4.11 of Chapter II and 4.3 of Chapter IV).

Chapter I develops the necessary amount of universal algebra. Chapters II and III respectively construct and analyse a model of the Propositional Calculus, introducing in simple form many of the ideas needed for the more complex First-Order Predicate Calculus, which is studied in Chapter IV. In Chapter V, we consider first-order mathematical theories, i.e., theories built on the First-Order Predicate Calculus, thus building models of parts of mathematics. As set theory is usually regarded as the basis on which the rest of mathematics is constructed, we devote Chapter VI to a study of first-order Zermelo-Fraenkel Set Theory. Chapter VII, on Ultraproducts, discusses a technique for constructing new models of a theory from a given collection of models. Chapter VIII, which is an introduction to Non-Standard Analysis, is included as an example of mathematical logic assisting in the study of another branch of mathematics. Decision processes are investigated in Chapter IX, and we prove there the non-existence of decision processes for a number of problems. In Chapter X, we discuss two decision problems from other branches of mathematics and indicate how the results of Chapter IX may be applied.

This book is intended to make mathematical logic available to mathematicians working in other branches of mathematics. We have included what we consider to be the essential basic theory, some useful techniques, and some indications of ways in which the theory might be of use in other branches of mathematics.

We have included a number of exercises. Some of these fill in minor gaps in our exposition of the section in which they appear. Others indicate aspects of the subject which have been ignored in the text. Some are to help in understanding the text by applying ideas and methods to special cases. Occasionally, an exercise asks for the construction of a FORTRAN program. In such cases, the solution should be based on integer arithmetic, and not depend on any special logical properties of FORTRAN or of any other programming language.

The layout of the text is as follows. Each chapter is divided into numbered sections, and definitions, theorems, exercises, etc. are numbered consecutively within each section. For example, the number 2.4 refers to the fourth item in the second section of the current chapter. A reference to an item in some other chapter always includes the chapter number in addition to item and section numbers.

We thank the many mathematical colleagues, particularly Paul Halmos and Peter Hilton, who encouraged and advised us in this project. We are especially indebted to Gordon Monro for suggesting many improvements and for providing many exercises. We thank Mrs. Blakestone and Miss Kicinski for the excellent typescript they produced.

Donald W. Barnes, John M. Mack

Table of Contents

Preface v

Chapter I Universal Algebra 1

§1 Introduction 1
§2 Free Algebras 4
§3 Varieties of Algebras 7
§4 Relatively Free Algebras 8

Chapter II Propositional Calculus 11

§1 Introduction 11
§2 Algebras of Propositions 11
§3 Truth in the Propositional Calculus. 13
§4 Proof in the Propositional Calculus. 14

Chapter III Properties of the Propositional Calculus . . . 18

§1 Introduction 18
§2 Soundness and Adequacy of Prop(X) 19
§3 Truth Functions and Decidability for Prop(X) . . . 22

Chapter IV Predicate Calculus 26

§1 Algebras of Predicates 26
§2 Interpretations 29
§3 Proof in Pred(V, $\mathscr{R}$) 30
§4 Properties of Pred(V, $\mathscr{R}$) 32

Chapter V First-Order Mathematics 38

§1 Predicate Calculus with Identity 38
§2 First-Order Mathematical Theories 39
§3 Properties of First-Order Theories 43
§4 Reduction of Quantifiers 48

Chapter VI Zermelo-Fraenkel Set Theory 52

§1 Introduction 52
§2 The Axioms of ZF 52
§3 First-Order ZF 56
§4 The Peano Axioms 58

Chapter VII Ultraproducts **62**

§1 Ultraproducts 62
§2 Non-Principal Ultrafilters 64
§3 The Existence of an Algebraic Closure 66
§4 Non-Trivial Ultrapowers 67
§5 Ultrapowers of Number Systems 68
§6 Direct Limits 70

Chapter VIII Non-Standard Models **74**

§1 Elementary Standard Systems. 74
§2 Reduction of the Order 75
§3 Enlargements 76
§4 Standard Relations 78
§5 Internal Relations. 79
§6 Non-Standard Analysis 80

Chapter IX Turing Machines and Gödel Numbers **85**

§1 Decision Processes 85
§2 Turing Machines 85
§3 Recursive Functions 89
§4 Gödel Numbers 90
§5 Insoluble Problems in Mathematics. 93
§6 Insoluble Problems in Arithmetic 96
§7 Undecidability of the Predicate Calculus 101

Chapter X Hilbert's Tenth Problem, Word Problems . . . **105**

§1 Hilbert's Tenth Problem 105
§2 Word Problems 110

References and Further Reading 115

Index of Notations 117

Subject Index 119

An Algebraic Introduction to Mathematical Logic

Chapter I

Universal Algebra

§1 Introduction

The reader will be familiar with the presentation and study of various algebraic systems (for example, groups, rings, modules) as axiomatic systems consisting of sets with certain operations satisfying certain conditions. The reader will also be aware that ideas and theorems, useful for the study of one type of system, can frequently be adapted to other related systems by making the obvious necessary modifications.

In this book we shall study and use a number of systems whose types are related, but which are possibly unfamiliar to the reader. Hence there is obvious advantage in beginning with the study of a single axiomatic theory which includes as special cases all the systems we shall use. This theory is known as universal algebra, and it deals with systems having arbitrary sets of operations. We shall want to avoid, as far as possible, axioms asserting the existence of elements with special properties (for example, the identity element in group theory), preferring the axioms satisfied by operations to take the form of equations, and we shall be able to achieve this by giving a sufficiently broad definition of "operation". We first recall some elementary facts.

An n-ary relation ρ on the sets $A_1, \ldots, A_n$ is specified by giving those ordered n-tuples $(a_1, \ldots, a_n)$ of elements $a_i \in A_i$ which are in the relation ρ. Thus such a relation is specified by giving those elements $(a_1, \ldots, a_n)$ of the product set $A_1 \times \cdots \times A_n$ which are in ρ, and hence an n-ary relation on $A_1, \ldots, A_n$ is simply a subset of $A_1 \times \cdots \times A_n$. For binary relations, the notation "$a_1 \rho a_2$" is commonly used to express "(a_1, a_2) is in the relation ρ", but we shall usually write this as either "$(a_1, a_2) \in \rho$" or "$\rho(a_1, a_2)$", because each of these notations extends naturally to n-ary relations for any n.

A function $f: A \to B$ is a binary relation on A and B such that, for each $a \in A$, there is exactly one $b \in B$ for which $(a, b) \in f$. It is usual to write this as $f(a) = b$. A function $f(x, y)$ "of two variables" $x \in A, y \in B$, with values in C, is simply a function $f: A \times B \to C$. For each $a \in A$ and $b \in B$, $(a, b) \in A \times B$ and $f((a, b)) \in C$. It is of course usual to omit one set of brackets. There are advantages in retaining the variables x, y in the function notation. Later in this chapter, we will discuss what is meant by variables and give a definition which will justify their use.

Preliminary Definition of Operation. *An n-ary operation on the set A is a function $t: A^n \to A$. The number n is called the arity of t.*

Examples

1.1. Multiplication in a group is a binary operation. The $*$-product of two elements a, b is written $a*b$ or simply ab instead of the more systematic $*(a, b)$.

1.2. In a group G, we can define a unary operation $i: G \to G$ by putting $i(a) = a^{-1}$.

1.3. A 0-ary operation on a set A is a function from the set A^0 (whose only element is the empty set $\varnothing$) to the set A, and hence can be regarded as a distinguished element of A. Such an operation arises naturally in group theory, where the 0-ary operation e gives the identity element of the group G.

One often considers several different groups in group theory. If G, H are groups, each has its multiplication operation: $*_G: G \times G \to G$ and $*_H: H \times H \to H$, but one rarely uses distinctive notations for the two multiplications. In practice, the same notation $*$ is used for both, and in fact multiplication is regarded as an operation defined for all groups. The definition of operation given above is clearly not adequate for this usage of the word.

Here is another example demonstrating that our preliminary definition of operation does not match common usage. A ring R is usually defined as a set R with two binary operations $+$, $\times$ satisfying certain axioms. A commonly occurring example of a ring is the zero ring where $R = \{0\}$. In this case, there is only one function $R \times R \to R$, and so $+$, $\times$ are the same function, even though $+$ and $\times$ are still considered distinct operations.

We now give a series of definitions which will overcome the objections raised above.

Definition 1.4. A *type* $\mathscr{T}$ is a set T together with a function $\mathrm{ar}: T \to \mathrm{N}$, from T into the non-negative integers. We shall write $\mathscr{T} = (T, \mathrm{ar})$, or, more simply, abuse notation and denote the type by T. It is also convenient to denote by T_n the set $\{t \in T | \mathrm{ar}(t) = n\}$.

Definition 1.5. An *algebra A of type T*, or a *T-algebra*, is a set A together with, for each $t \in T$, a function $t_A: A^{\mathrm{ar}(t)} \to A$. The elements $t \in T_n$ are called n-ary T-algebra operations.

Observe that each t_A is an operation on the set A in the sense of our preliminary definition of operation. As is usual, we shall write simply $t(a_1, \ldots, a_n)$ for the element $t_A(a_1, \ldots, a_n)$, and we shall denote the algebra by the same symbol A as is used to denote its set of elements.

Examples

1.6. Rings may be considered as algebras of type $T = (\{0, -, +, \cdot\}, \mathrm{ar})$, where $\mathrm{ar}(0) = 0, \mathrm{ar}(-) = 1, \mathrm{ar}(+) = 2, \mathrm{ar}(\cdot) = 2$. We do not claim that such T-algebras are necessarily rings, we simply assert that each ring is an example of a T-algebra for the T given above.

1.7. If R is a given ring, then a module over R may be regarded as a particular example of a T-algebra of type $T = (\{0, -, +\} \cup R, \text{ar})$, where $\text{ar}(0) = 0$, $\text{ar}(-) = 1$, $\text{ar}(+) = 2$, and $\text{ar}(\lambda) = 1$ for each $\lambda \in R$. The first three operations specify the group structure of the module, while the remaining operations correspond to the action of the ring elements.

1.8. Let S be a given ring. Rings R which contain S as subring may be considered as T-algebras, where $T = (\{0, -, +, \cdot\} \cup S, \text{ar})$, $\text{ar}(0) = 0$, $\text{ar}(-) = 1$, $\text{ar}(+) = 2$, $\text{ar}(\cdot) = 2$, and $\text{ar}(s) = 0$ for each $s \in S$. The effect of the S-operations is to distinguish certain elements of R.

Definition 1.9. T-algebras A, B are *equal* if and only if $A = B$ and $t_A = t_B$ for all $t \in T$.

Exercise 1.10. Give an example of unequal T-algebras on the same set A.

Definition 1.11. If A is a T-algebra, a subset B of A is called a *T-subalgebra* of A if it forms a T-algebra with operations the restrictions to B of those on A, i.e., if for all n and for all $t \in T_n$ and $b_1, \ldots, b_n \in B$, we have $t_A(b_1, \ldots, b_n) \in B$.

Any intersection of subalgebras is a subalgebra, and so, given any subset X of A, there is a unique smallest subalgebra containing X—namely, the subalgebra $\cap\{U | U \text{ subalgebra of } A,\ U \supseteq X\}$. We call this the subalgebra generated by X and denote it by $\langle X \rangle_T$, or if there is no risk of confusion, by $\langle X \rangle$.

Exercises

1.12. A is a T-algebra. Show that $\varnothing$ is a subalgebra if and only if $T_0 = \varnothing$. Show that for all T, every T-algebra has a unique smallest subalgebra.

'Many familiar algebraic systems may be regarded as T-algebras for more than one choice of T. However, the subsets which form T-subalgebras may well depend on the choice of T.

1.13. Groups may be regarded as special cases of T-algebras where $T = (\{*\}, \text{ar})$ with $\text{ar}(*) = 2$, or of T'-algebras, where $T' = (\{e, i, *\}, \text{ar})$, $\text{ar}(e) = 0$, $\text{ar}(i) = 1$, $\text{ar}(*) = 2$. Show that every T'-subalgebra of a group is a subgroup, but that not every non-empty T-subalgebra need be a group. Show that if G is a finite group, then every non-empty T-subalgebra of G is itself a group.

Definition 1.14. Let A, B be T-algebras. A *homomorphism* of A into B is a function $\varphi: A \to B$ such that, for all $t \in T$ and all $a_1, \ldots, a_n \in A$ ($n = \text{ar}(t)$), we have

$$\varphi(t_A(a_1, \ldots, a_n)) = t_B(\varphi(a_1), \ldots, \varphi(a_n)).$$

This condition is often expressed as "*φ preserves all the operations of T*".

Clearly, the composition of two homomorphisms is a homomorphism. Further, if $\varphi : A \to B$ is a homomorphism and is invertible, then the inverse function $\varphi^{-1} : B \to A$ is also a homomorphism. In this case we call φ an *isomorphism* and say that A and B are *isomorphic*.

§2 Free Algebras

Definition 2.1. Let X be any set, let F be a T-algebra and let $\sigma : X \to F$ be a function. We say that F (more strictly (F, σ)) is a *free T-algebra* on the set X of *free generators* if, for every T-algebra A and function $\tau : X \to A$, there exists a unique homomorphism $\varphi : F \to A$ such that $\varphi\sigma = \tau$:

$$\begin{tikzcd} X \arrow[rr, "\sigma"] \arrow[dr, "\tau"'] & & F \arrow[dl, dashed, "\varphi"] \\ & A & \end{tikzcd}$$

Observe that if (F, σ) is free, then σ is injective. For it is easily seen that there exists a T-algebra with more than one element, and hence if x_1, x_2 are distinct elements of X, then for some A and τ we have $\tau(x_1) \neq \tau(x_2)$, which implies $\sigma(x_1) \neq \sigma(x_2)$.

The next theorem asserts the existence of a free T-algebra on a set X, and the proof is constructive. Informally, one could describe the free T-algebra on X as the collection of all formal expressions that can be formed from X and T by using only finitely many elements of X and T in any one expression. But to say precisely what is meant by a formal expression in the elements of X using the operations of T is tantamount to constructing the free algebra.

Theorem 2.2. *For any set X and any type T, there exists a free T-algebra on X. This free T-algebra on X is unique up to isomorphism.*

Proof. (a) *Uniqueness.* We show first that if (F, σ) is free on X, and if $\varphi : F \to F$ is a homomorphism such that $\varphi\sigma = \sigma$, then $\varphi = 1_F$, the identity map on F. To show this, we take $A = F$ and $\tau = \sigma$ in the defining condition. Then $1_F : F \to F$ has the required property for φ, and hence by its uniqueness is the only such map.

Now let (F, σ) and (F', σ') be free on X.

$$\begin{tikzcd}
X \arrow[rr, "\sigma"] \arrow[dr, "\sigma'"'] & & F \arrow[dl, dashed, shift right, "\varphi"'] \\
& F' \arrow[ur, dashed, shift right, "\varphi'"'] &
\end{tikzcd}$$

Since (F, σ) is free, there exists a homomorphism $\varphi: F \to F'$ such that $\varphi\sigma = \sigma'$. Since (F', σ') is free, there exists a homomorphism $\varphi': F' \to F$ such that $\varphi'\sigma' = \sigma$. Hence $\varphi'\varphi\sigma = \varphi'\sigma' = \sigma$, and by the result above, $\varphi'\varphi = 1_F$. Similarly, $\varphi\varphi' = 1_{F'}$. Thus φ, φ' are mutually inverse isomorphisms, and so uniqueness is proved.

(b) *Existence.* An algebra F will be constructed as a union of sets F_n $(n \in \mathrm{N})$, which are defined inductively as follows.

(i) F_0 is the disjoint union of X and T_0.

(ii) Assume F_r is defined for $0 \leqslant r < n$. Then define

$$F_n = \left\{(t, a_1, \ldots, a_k) \middle| t \in T, \operatorname{ar}(t) = k, a_i \in F_{r_i}, \sum_{i=1}^{k} r_i = n - 1\right\}.$$

(iii) Put $F = \bigcup_{n \in \mathrm{N}} F_n$.

The set F is now given. To make it into a T-algebra, we must specify the action of the operations $t \in T$.

(iv) If $t \in T_k$ and $a_1, \ldots, a_k \in F$, put $t(a_1, \ldots, a_k) = (t, a_1, \ldots, a_k)$. In particular, if $t \in T_0$, then t_F is the element t of F_0.

This makes F into a T-algebra. To complete the construction, we must give the map $\sigma: X \to F$.

(v) For each $x \in X$, put $\sigma(x) = x \in F_0$.

Finally, we have to prove that F is free on X, i.e., we must show that if A is any T-algebra and $\tau: X \to A$ any map of X into A, then there exists a unique homomorphism $\varphi: F \to A$ such that $\varphi\sigma = \tau$. We do this by constructing inductively the restriction φ_n of φ to F_n and by showing that φ_n is completely determined by τ and the φ_k for $k < n$.

We have $F_0 = T_0 \cup X$. The homomorphism condition requires $\varphi_0(t_F) = t_A$ for $t \in T_0$, while for $x \in X$ we require $\varphi\sigma(x) = \tau(x)$, and so we must have

$\varphi_0(x) = \tau(x)$. Thus $\varphi_0 : F_0 \to A$ is defined, and is uniquely determined by the conditions to be satisfied by φ.

Suppose that φ_k is defined and uniquely determined for $k < n$. An element of F_n $(n > 0)$ is of the form $(t, a_1 \ldots, a_k)$, where $t \in T_k$, $a_i \in F_{r_i}$ and $\sum_{i=1}^{k} r_i = n - 1$. Thus $\varphi_{r_i}(a_i)$ is already uniquely defined for $i = 1, \ldots, k$. Furthermore, since $(t, a_1, \ldots, a_k) = t(a_1, \ldots, a_k)$, and since the homomorphism property of φ requires that

$$\varphi(t, a_1, \ldots, a_k) = t(\varphi(a_1), \ldots, \varphi(a_k)),$$

we must define

$$\varphi_n(t, a_1, \ldots, a_k) = t(\varphi_{r_1}(a_1), \ldots, \varphi_{r_k}(a_k)).$$

This determines φ_n uniquely, and as each element of F belongs to exactly one subset F_n, on putting $\varphi(\alpha) = \varphi_n(\alpha)$ for $\alpha \in F_n$ $(n \geqslant 0)$, we see that φ is a homomorphism from F to A satisfying $\varphi\sigma(x) = \varphi_0(x) = \tau(x)$ for all $x \in X$ as required, and that φ is the only such homomorphism. □

The above inductive construction of the free T-algebra F fits in with its informal description—each F_n is a collection of "T-expressions", increasing in complexity with n. The notion of a T-expression is useful for an arbitrary T-algebra, so we shall formalise it, making use of free T-algebras to do so.

Let A be any T-algebra, and let F be the free T-algebra on the set $X_n = \{x_1, \ldots, x_n\}$. For any (not necessarily distinct) elements $a_1, \ldots, a_n \in A$, there exists a unique homomorphism $\varphi : F \to A$ with $\varphi(x_i) = a_i$ $(i = 1, \ldots, n)$. If $w \in F$, then $\varphi(w)$ is an element of A which is uniquely determined by $a_1, \ldots, a_n$. Hence we may define a function $w_A : A^n \to A$ by putting $w_A(a_1, \ldots, a_n) = \varphi(w)$. We omit the subscript A and write simply $w(a_1, \ldots, a_n)$. If in particular we take $A = F$ and $a_i = x_i$ $(i = 1, \ldots, n)$, then φ is the identity and $w(x_1, \ldots, x_n) = w$.

Definition 2.3. A *T-word* in the *variables* $x_1, \ldots, x_n$ is an element of the free T-algebra on the set $X_n = \{x_1, \ldots, x_n\}$ of free generators.

Definition 2.4. A *word* in the elements $a_1, \ldots, a_n$ of a T-algebra A is an element $w(a_1, \ldots, a_n) \in A$, where w is a T-word in the variables $x_1, \ldots, x_n$.

We have used and even implicitly defined the term "variable" in the above definitions. In normal usage, a variable is "defined" as a symbol for which any element of the appropriate kind may be substituted. We give a formal definition of variable, confirming that our variables have this usual property.

Definition 2.5. A *T-algebra variable* is an element of the free generating set of a free T-algebra.

Among the words in the variables $x_1, \ldots, x_n$ are the words x_i $(i = 1, \ldots, n)$, having the property that $x_i(a_1, \ldots, a_n) = a_i$. Thus variables may also be regarded as coordinate functions. The concept of a coordinate function certainly provides the most convenient definition of variable for use in analysis. For example, when we speak of a function $f(x, y)$ as a function of two real variables x, y, we have a function f, defined on some subset of $\mathbf{R} \times \mathbf{R}$, together with coordinate projections $x(a, b) = a$, $y(a, b) = b$ $(a, b \in \mathbf{R})$, and $f(x, y)$ is in fact the composite function $f(a, b) = f(x(a, b), y(a, b))$.

Exercises

2.6. T consists of one unary operation, and F is the free T-algebra on a one-element set X. How many elements are there in F_n? How many elements are there in F?

2.7. If T is empty and X is any set, show that X is the free T-algebra on X.

2.8. T consists of a single binary operation, and F is the free T-algebra on a one-element set X. How many elements are there in F?

2.9. If T consists of one 0-ary operation and one 2-ary operation, and if $X = \emptyset$, then the free T-algebra F on X is countable.

2.10. T is finite or countable, and contains at least one 0-ary operation and at least one operation t with $\operatorname{ar}(t) > 0$. X is finite or countable. Prove that F is countable.

§3 Varieties of Algebras

Let F be the free T-algebra on the countable set $X = \{x_1, x_2, \ldots\}$ of variables. Although each element of F is a word in some finite subset $X_n = \{x_1, \ldots, x_n\}$, we shall consider sets of words for which there may be no bound to the number of variables in the words.

Definition 3.1. An *identical relation* on T-algebras is a pair (u, v) of elements of F.

There is an n for which u, v are in the free algebra on X_n, and we say that (u, v) is an *n-variable identical relation* for any such n.

Definition 3.2. The T-algebra A *satisfies* the n-variable identical relation (u, v), or (u, v) is a *law* of A, if $u(a_1, \ldots, a_n) = v(a_1, \ldots, a_n)$ for all $a_1, \ldots, a_n \in A$.

Equivalently, (u, v) is a law of A if $\varphi(u) = \varphi(v)$ for every homomorphism $\varphi: F \to A$.

Definition 3.3. Let L be a set of identical relations on T-algebras. The class V of all T-algebras which satisfy all the identical relations in L is called the *variety of T-algebras defined by L*. The *laws of the variety* are all the identical relations satisfied by every algebra of V.

Note that the set of laws of the variety includes L, but may be larger.

Examples

3.4. T consists of a single binary operation $*$, and L has the one element $(x_1*(x_2*x_3), (x_1*x_2)*x_3)$. If A satisfies this identical relation, then $a*(b*c) = (a*b)*c$ for all $a, b, c \in A$. Thus the operation on A is associative and A is a semigroup. The variety defined by L in this case is the class of all semigroups.

3.5. T consists of 0-ary, 1-ary and 2-ary operations e, i, $*$ respectively. L has the three elements

$$\begin{gathered}(x_1*(x_2*x_3), (x_1*x_2)*x_3),\\ (e*x_1, x_1),\\ (i(x_1)*x_1, e).\end{gathered}$$

The first law ensures that $*$ is an associative operation in every algebra of the variety defined by L. The second shows that the distinguished element e is always a left identity, while the third guarantees that $i(a)$ is a left inverse of the element a. Hence the algebras of the variety are groups.

Exercises

3.6. Show that the class of all abelian groups is a variety.

3.7. R is a ring with 1. Show that the class of unital left R-modules is a variety.

3.8. S is a commutative ring with 1. Show that the class of commutative rings R with $1_R = 1_S$ and which contain S as a subring is a variety.

3.9. Is the class of finite groups a variety?

§4 Relatively Free Algebras

Let V be the variety of T-algebras defined by the set L of laws.

Definition 4.1. A T-algebra R in the variety V is the (*relatively*) *free algebra of* V on the set X of (*relatively*) *free generators* (where a function $\sigma: X \to R$ is given, usually as an inclusion) if, for every algebra A in V and every function $\tau: X \to A$, there exists a unique homomorphism $\varphi: R \to A$ such that $\varphi\sigma = \tau$.

This definition differs from the earlier definition of a free algebra only in that we consider here only algebras in V.

Definition 4.2. An algebra is *relatively free* if it is a free algebra of some variety.

Theorem 4.3. *For any type T, and any set L of laws, let V be the variety of T-algebras defined by L. For any set X, there exists a free T-algebra of V on X.*

Proof: Let (F, ρ) be the free T-algebra on X. A congruence relation on F is defined by putting $u \sim v$ (where $u, v \in F$) if $\varphi(u) = \varphi(v)$ for every homomorphism φ of F into an algebra in V. Clearly $\sim$ is an equivalence relation on F. If now $t \in T_k$ and $u_i \sim v_i$ $(i = 1, \dots, k)$, then for every such homomorphism φ, $\varphi(u_i) = \varphi(v_i)$, and so

$$\varphi(t(u_1, \dots, u_k)) = t(\varphi(u_1), \dots, \varphi(u_k)) = t(\varphi(v_1), \dots, \varphi(v_k)) = \varphi(t(v_1, \dots, v_k)),$$

verifying that φ is a congruence relation.

We define R to be the set of congruence classes of elements of F with respect to this congruence relation. Denoting the congruence class containing u by $\bar{u}$, we define the action of $t \in T_k$ on R by putting $t(\bar{u}_1, \dots, \bar{u}_k) = \overline{t(u_1, \dots, u_k)}$. This definition is independent of the choice of representatives $u_1, \dots, u_k$ of the classes $\bar{u}_1, \dots, \bar{u}_k$, and makes R a T-algebra. Also, the map $u \to \bar{u}$ is clearly a homomorphism $\eta : F \to R$. Finally, we define $\sigma : X \to R$ by $\sigma(x) = \overline{\rho(x)}$.

We now prove that (R, σ) is relatively free on X. Let A be any algebra in V, and let $\tau : X \to A$ be any function from X into A. Because (F, ρ) is free, there exists a unique homomorphism $\psi : F \to A$ such that $\psi\rho = \tau$.

$$\begin{tikzcd}
X \arrow[rr, "\rho"] \arrow[dr, "\sigma"] \arrow[ddr, "\tau"'] & & F \arrow[dl, "\eta"'] \arrow[ddl, dashed, "\psi"] \\
& R \arrow[d, dashed, "\varphi"] & \\
& A &
\end{tikzcd}$$

For $\bar{u} \in R$, we define $\varphi(\bar{u}) = \psi(u)$. This is independent of the choice of representative u of the element $\bar{u}$, since if $\bar{u} = \bar{v}$, then $\psi(u) = \psi(v)$. The map $\varphi : R \to A$ is clearly a homomorphism, and $\varphi\sigma = \varphi\eta\rho = \psi\rho = \tau$. If $\varphi' : R \to A$ is another homomorphism such that $\varphi'\sigma = \tau$, then $\varphi'\eta\rho = \tau$ and therefore $\varphi'\eta = \psi$. Consequently for each element $\bar{u} \in R$ we have

$$\varphi'(\bar{u}) = \varphi'\eta(u) = \psi(u) = \varphi(\bar{u}),$$

and hence $\varphi' = \varphi$. □

When considering only the algebras of a given variety V, we may redefine variables and words accordingly. Thus we define a *V-variable* as an element of the free generating set of a free algebra of V, and a *V-word* in the V-variables $x_1, \dots, x_n$ as an element of the free algebra of V on the free generators $\{x_1, \dots, x_n\}$.

Examples

4.4. T consists of a single binary operation which we shall write as juxtaposition. Let V be the variety of associative T-algebras. Then all products in the free T-algebra obtained by any bracketing of $x_1, \ldots, x_n$, taken in that order, are congruent under the congruence relation used in our construction of the relatively free algebra, and correspond to the one word $x_1 x_2 \cdots x_n$ of V. We observe that in this example, all elements of the absolutely free algebra F, which map to a given element $x_1 x_2 \cdots x_n$ of the relatively free algebra, come from the same layer F_{n-1} of F.

4.5. T consists of a 0-ary, a 1-ary and a 2-ary operation. V is the variety of abelian groups, defined by the laws given in Example 3.5 together with the law $(x_1 x_2, x_2 x_1)$. In this case, the relatively free algebra on $\{x_1, \ldots, x_n\}$ is the set of all $x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$ (or equivalently the set of all n-tuples $(r_1, \ldots, r_n)$) with $r_i \in \mathbf{Z}$. Here the layer property of Example 4.4 does not hold, because, for example, we have the identity $e \in F_0$, $x_1^{-1} \in F_1$, $x_1^{-1} * x_1 \in F_2$ and yet $\bar{e} = \overline{x_1^{-1} * x_1}$.

Exercises

4.6. K is a field. Show that vector spaces over K form a variety V of algebras, and that every vector space over K is a free algebra of V.

4.7. R is a commutative ring with 1 and V is the variety of commutative rings S which contain R as a subring and in which 1_R is a multiplicative identity of S. Show that the free algebra of V on the set X of variables is the polynomial ring over R in the elements of X.

Chapter II

Propositional Calculus

§1 Introduction

Mathematical logic is the study of logic as a mathematical theory. Following the usual procedure of applied mathematics, we construct a mathematical model of the system to be studied, and then conduct what is essentially a pure mathematical investigation of the properties of our model. Since this book is intended for mathematicians, the system we propose to study is not general logic but the logic used in mathematics. By this restriction, we achieve considerable simplification, because we do not have to worry about precise meanings of words—in mathematics, words have precisely defined meanings. Furthermore, we are free of reasoning based on things such as emotive argument, which must be accounted for in any theory of general logic. Finally, the nature of the real world need not concern us, since the world we shall study is the purely conceptual one of pure mathematics.

In any formal study of logic, the language and system of reasoning needed to carry out the investigation is called the meta-language or meta-logic. As we are constructing a mathematical model of logic, our meta-language is mathematics, and so all our existing knowledge of mathematics is available for possible application to our model. We shall make specific use of informal set theory (including cardinal numbers and Zorn's lemma) and of the universal algebra developed in Chapter I.

For the purpose of our study, it suffices to describe mathematics as consisting of assertions that if certain statements are true then so are certain other statements, and of arguments justifying these assertions. Hence a model of mathematical reasoning must include a set of objects which we call statements or propositions, some concept of truth, and some concept of a proof. Once a model is constructed, the main subject of investigation is the relationship between truth and proof. We shall begin by constructing a model of the simpler parts of mathematical reasoning. This model is called the Propositional Calculus. Later, we shall construct a more refined model (known as the First-Order Predicate Calculus), copying more complicated parts of the reasoning used in mathematics.

§2 Algebras of Propositions

The Propositional Calculus considers ways in which simple statements may be combined to form more complex statements, and studies how the truth or falsity of complex statements is related to that of their component statements. Some of the ways in which statements are combined in mathematics are as follows. We often use "and" to combine statements, and we write $p \wedge q$ for the statement "p and q", which is regarded as true if and only if both the statements p, q are true. We frequently assert that (at least) one of two possibilities is true, and we write $p \vee q$ for the statement "p or q", which we consider to be true if at least one of p, q is true and false if both p and q are false. We often assert that some statement is false, and we write $\sim p$ (read "not p") for the statement "p is false", which is regarded as true if and only if p is false. Another common way of linking two statements is through an assertion "if p is true, then so is q". For this we write "$p \Rightarrow q$" (read "p implies q"), which, in mathematical usage, is true unless q is false and p is true.

We want our simple model to imitate the above constructions, so we want our set of propositions to be an algebra with respect to the four operations given above. This could be done by taking the free algebra with these operations, but we know that in ordinary usage, the four operations are not independent. Thus a simpler system is suggested, in which we choose some basic operations which will enable us to define all the above operations. This may be done in many ways, some of which are explored in exercises at the end of Chapter III, where they may be studied more thoroughly. We choose a way which is perhaps not the natural one, but which has advantages in that it simplifies the development of the theory. Our choice rests on the fact that in mathematics, a result is often proved by showing that the denial of the result leads to a contradiction. We introduce into our notation a symbol for a contradiction by specifying that our algebra will have a distinguished element (i.e., a 0-ary operation) F, which we will think of as a contradiction or falsehood.

Definition 2.1. Let $T = \{F, \Rightarrow\}$, where F is a 0-ary operation and $\Rightarrow$ is a binary operation. Any T-algebra is called a *proposition algebra.*

Definition 2.2. The *proposition algebra $P(X)$ of the propositional calculus on the set X of propositional variables* is the free T-algebra on X.

Example 2.3. The algebra $\mathbf{Z}_2$ of integers mod 2 can be made into a proposition algebra by defining $F_{\mathbf{Z}_2} = 0$ and $m \Rightarrow n = 1 + m(1 + n)$. We shall make frequent use of this example.

In any proposition algebra, we introduce the further operations $\sim$, $\vee$, $\wedge$, $\Leftrightarrow$ by defining

$$\begin{aligned} \sim p &= p \Rightarrow F \\ p \vee q &= (\sim p) \Rightarrow q \\ p \wedge q &= \sim(\sim p \vee \sim q) \\ p \Leftrightarrow q &= (p \Rightarrow q) \wedge (q \Rightarrow p). \end{aligned}$$

We point out that the above are not statements in our proposition algebras, because the symbol $=$ is not an operation in our proposition algebras. The first equation says that $\sim p$ is a notation for the element $p \Rightarrow F$ of our algebra. We shall often omit brackets, as we did above in writing $\sim p \vee \sim q$ for $(\sim p) \vee (\sim q)$.

Exercises

2.4. Show that our definitions of $\sim$, $\vee$, $\wedge$, $\Leftrightarrow$ conform to normal usage.
2.5. Express $\sim$, $\vee$, $\wedge$ in $\mathbf{Z}_2$ in terms of multiplication and addition.
2.6. Is $\mathbf{Z}_2$ a free proposition algebra?

§3 Truth in the Propositional Calculus

Having determined the form of our algebra of propositions, we must now find a meaning for the concept of truth applied to our propositions. We are guided here by the observation that in ordinary mathematical usage, the truth or falsity of the compound statement $p \Rightarrow q$ is determined completely once the truth or falsity of each of p, q is specified. Every simple statement is given a value—true or false—and the truth or falsity of any compound statement depends on and is determined by the truth values of its components. This leads us to consider valuations on $P(X)$, i.e., functions which assign to each element $p \in P(X)$ one of two possible values, which for convenience are denoted by 0, 1. We are then considering functions $v: P(X) \rightarrow \mathbf{Z}_2$, interpreting $v(p) = 1$ as meaning "p is true", and $v(p) = 0$ as "p is false". In order that a valuation act properly on compound propositions, the functions v must be proposition algebra homomorphisms.

Definition 3.1. A *valuation* of $P(X)$ is a proposition algebra homomorphism $v: P(X) \rightarrow \mathbf{Z}_2$. We say that $p \in P(X)$ is *true with respect to* v if $v(p) = 1$, and that p is *false with respect to* v if $v(p) = 0$.

Since X is a set of free generators of $P(X)$, the values $v(x)$ for $x \in X$ may be assigned arbitrarily. These values, once assigned, determine the homomorphism v uniquely and so determine $v(p)$ for all $p \in P(X)$.

In ordinary usage, the interesting and important notion relating the truth values of statements is that of consequence—a statement q is a consequence of statements $p_1, \ldots, p_n$ if q is true of every mathematical system in which $p_1, \ldots, p_n$ are all true. This idea is incorporated in our model by considering valuations which assign the value 1 to all of $p_1, \ldots, p_n$.

Definition 3.2. Let $A \subseteq P(X)$ and $q \in P(X)$. We say that q is a *consequence* of the set A of assumptions, or that A *semantically implies* q, if $v(q) = 1$ for every valuation v such that $v(p) = 1$ for all $p \in A$. We shall write this $A \vDash q$, and we shall denote by $\operatorname{Con}(A)$ the set $\{p \in P(X) | A \vDash p\}$ of all consequences of A.

Definition 3.3. Let $p \in P(X)$. We say that p is *valid*, or is a *tautology*, if $v(p) = 1$ for every valuation v of $P(X)$.

Thus p is a tautology if $\varnothing \models p$. We shall write this simply as $\models p$. Note that $A \models p$ is not a proposition (i.e., not an element of $P(X)$), but simply a statement in the meta-language about our model.

Examples

3.4. $\{q\} \models p \Rightarrow q$. For if v is any valuation with $v(q) = 1$, then

$$v(p \Rightarrow q) = v(p) \Rightarrow v(q) = v(p) \Rightarrow 1 = 1 + v(p)(1 + 1) = 1.$$

3.5. $\models p \Rightarrow p$. For if v is any valuation, then

$$v(p \Rightarrow p) = v(p) \Rightarrow v(p) = 1 + v(p)(1 + v(p)) = 1,$$

since $x(1 + x) = 0$ for all $x \in \mathbf{Z}_2$.

Exercises

3.6. Show that $\{F\} \models p$ for all $p \in P(X)$.

3.7. Show that $\{p, p \Rightarrow q\} \models q$ and $\{p, \sim q \Rightarrow \sim p\} \models q$ for all $p, q \in P(X)$.

3.8. Show that $p \Rightarrow (q \Rightarrow p)$, $(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$ and $\sim \sim p \Rightarrow p$ are tautologies, for all $p, q, r \in P(X)$.

Lemma 3.9. Con *is a closure operation on* $P(X)$, *that is, it has the properties*

(i) $A \subseteq \mathrm{Con}(A)$,
(ii) *If* $A_1 \subseteq A_2$, *then* $\mathrm{Con}(A_1) \subseteq \mathrm{Con}(A_2)$,
(iii) $\mathrm{Con}(\mathrm{Con}(A)) = \mathrm{Con}(A)$.

Proof:

(i) Trivial.
(ii) Suppose $q \in \mathrm{Con}(A_1)$. Let v be any valuation such that $v(A_2) \subseteq \{1\}$. Then $v(A_1) \subseteq \{1\}$ and so $v(q) = 1$ since $q \in \mathrm{Con}(A_1)$. Hence $q \in \mathrm{Con}(A_2)$.
(iii) Suppose $q \in \mathrm{Con}(\mathrm{Con}(A))$, and let v be a valuation such that $v(A) \subseteq \{1\}$. For all $p \in \mathrm{Con}(A)$, we have $v(p) = 1$ by the definition of $\mathrm{Con}(A)$. Thus $v(\mathrm{Con}(A)) \subseteq \{1\}$ and so $v(q) = 1$. Thus $q \in \mathrm{Con}(A)$. □

§4 Proof in the Propositional Calculus

A mathematical system is usually specified by certain statements called assumptions, which describe certain characteristic features of the system. A proof of some other property of the system consists of a succession of statements, ending in a statement of the desired property, in which each statement has been obtained from those before it in some acceptable manner. Apart from the particular assumptions of the system, which are considered acceptable at any step in a proof, we distinguish two methods which permit the addition of a statement to a given acceptable string of statements. There is a specific collection of statements which are considered acceptable additions in any mathematical proof—they can be regarded as underlying assumptions common to every mathematical system—and which we formalise as certain specified propositions which may be introduced at any stage into any proof. Such propositions are called the axioms of our model. The other permissible method consists of rules which specify, in terms of those statements already set down, particular statements which may be adduced. Rules of this kind, when formalised, are called the rules of inference of our model.

For the propositional calculus on the set X, we take as axioms all elements of the subset $\mathscr{A} = \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$ of $P(X)$,

where $\mathscr{A}_1 = \{p \Rightarrow (q \Rightarrow p) | p, q \in P(X)\}$,

$$\mathscr{A}_2 = \{(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) | p, q, r \in P(X)\},$$

and

$$\mathscr{A}_3 = \{\sim\sim p \Rightarrow p | p \in P(X)\}.$$

As our one rule of inference, we take the rule known as modus ponens: from p and $p \Rightarrow q$, deduce q. We may now give a formal definition of a proof.

Definition 4.1. Let $q \in P(X)$ and let $A \subseteq P(X)$. In the propositional calculus on the set X, a *proof of q from the assumptions A* is a finite sequence $p_1, p_2, \ldots, p_n$ of elements $p_i \in P(X)$ such that $p_n = q$ and for each i, either $p_i \in \mathscr{A} \cup A$ or for some $j, k < i$, we have $p_k = (p_j \Rightarrow p_i)$.

Definition 4.2. Let $q \in P(X)$ and let $A \subseteq P(X)$. We say that q is a *deduction* from A, or q is *provable* from A, or that A *syntactically implies* q, if there exists a proof of q from A. We shall write this $A \vdash q$, and we shall denote by $\text{Ded}(A)$ the set of all deductions from A.

Definition 4.3. Let $p \in P(X)$. We say that p is a *theorem* of the propositional calculus on X if there exists a proof of p from $\varnothing$.

Thus p is a theorem if $\varnothing \vdash p$, which we write simply as $\vdash p$.

Lemma 4.4. (i) *If* $q \in \text{Ded}(A)$, *then* $q \in \text{Ded}(A')$ *for some finite subset* A' *of* A.

(ii) Ded *is a closure operation on* $P(X)$.

Proof: (i) This holds because a proof of q from A, being a finite sequence of elements of $P(X)$, can contain only finitely many members of A.

(ii) The first two requirements for a closure operation are obviously met by Ded. Suppose now that $q \in \text{Ded}(\text{Ded}(A))$. Then there exists a proof $p_1, \ldots, p_n$ of q from $\text{Ded}(A)$. In this proof, certain (perhaps none) of the p_i, say $p_{i_1}, \dots, p_{i_r}$ are in Ded(A). Let $p_{i_j,1}, p_{i_j,2}, \dots, p_{i_j,r_j}$ be a proof of p_{i_j} from A. Replace each of the p_{i_j} in $p_1, \dots, p_n$ by its proof $p_{i_j,1}, \dots, p_{i_j,r_j}$. The resulting sequence is a proof of q from A. □

Examples

4.5. $\vdash p \Rightarrow p$. For any $p \in P(X)$, the following sequence $p_1, \dots, p_5$ is a proof of $p \Rightarrow p$:

$$\begin{aligned}
p_1 &= p \Rightarrow ((p \Rightarrow p) \Rightarrow p), && (\mathscr{A}_1)\\
p_2 &= (p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)), && (\mathscr{A}_2)\\
p_3 &= (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p), && (p_2 = p_1 \Rightarrow p_3)\\
p_4 &= p \Rightarrow (p \Rightarrow p), && (\mathscr{A}_1)\\
p_5 &= p \Rightarrow p. && (p_3 = p_4 \Rightarrow p_5)
\end{aligned}$$

The proof is the sequence $p_1, \dots, p_5$. These have been written on successive lines for ease of reading. We have placed notes alongside each step to explain why it can be included at that stage of the proof, but these notes are not part of the proof.

4.6. $\{q\} \vdash p \Rightarrow q$. A proof of this is $q \Rightarrow (p \Rightarrow q), q, p \Rightarrow q$.

4.7. $\vdash F \Rightarrow q$. For any $q \in P(X)$, the following is a proof:

$$\begin{aligned}
p_1 &= (\sim \sim q \Rightarrow q) \Rightarrow (F \Rightarrow (\sim \sim q \Rightarrow q)), && (\mathscr{A}_1)\\
p_2 &= \sim \sim q \Rightarrow q, && (\mathscr{A}_3)\\
p_3 &= F \Rightarrow (\sim \sim q \Rightarrow q), && (p_1 = p_2 \Rightarrow p_3)\\
p_4 &= (F \Rightarrow (\sim \sim q \Rightarrow q)) \Rightarrow ((F \Rightarrow \sim \sim q) \Rightarrow (F \Rightarrow q)), && (\mathscr{A}_2)\\
p_5 &= (F \Rightarrow \sim \sim q) \Rightarrow (F \Rightarrow q), && (p_4 = p_3 \Rightarrow p_5)\\
p_6 &= F \Rightarrow (\sim q \Rightarrow F) = F \Rightarrow \sim \sim q, && (\mathscr{A}_1)\\
p_7 &= F \Rightarrow q. && (p_5 = p_6 \Rightarrow p_7)
\end{aligned}$$

4.8. $\vdash \sim p \Rightarrow (p \Rightarrow q)$. A proof of this is the sequence $p_1, \dots, p_7$ of Example 4.7, followed by

$$\begin{aligned}
p_8 &= (F \Rightarrow q) \Rightarrow (p \Rightarrow (F \Rightarrow q)), && (\mathscr{A}_1)\\
p_9 &= p \Rightarrow (F \Rightarrow q), && (p_8 = p_7 \Rightarrow p_9)\\
p_{10} &= (p \Rightarrow (F \Rightarrow q)) \Rightarrow ((p \Rightarrow F) \Rightarrow (p \Rightarrow q)), && (\mathscr{A}_2)\\
p_{11} &= (p \Rightarrow F) \Rightarrow (p \Rightarrow q) = \sim p \Rightarrow (p \Rightarrow q). && (p_{10} = p_9 \Rightarrow p_{11})
\end{aligned}$$

The length of the proof needed for such a trivial result as $\sim p \Rightarrow (p \Rightarrow q)$ may well alarm a reader familiar with mathematical theorems and proofs. Ordinary mathematical proofs are very much abbreviated. For example, (allegedly) obvious steps are usually omitted, and previously established results are quoted without proof. Such devices are not available to us, because of the very restrictive nature of our definition of proof in the propositional calculus. We could reduce the lengths of many proofs if we extended our definition to include further rules of inference or abbreviative rules, but by doing so, we would complicate our study of the relationship between truth and proof, which is the principal object of the theory. We remark that in order to show that $\sim p \Rightarrow (p \Rightarrow q)$ is a theorem of the propositional calculus, it suffices to argue as follows: we have $\vdash F \Rightarrow q$, and the sequence $p_7, \ldots, p_{11}$ is a proof of $\sim p \Rightarrow (p \Rightarrow q)$ from the assumption $\{F \Rightarrow q\}$. Thus

$$\sim p \Rightarrow (p \Rightarrow q) \in \mathrm{Ded}(\{F \Rightarrow q\}) \subseteq \mathrm{Ded}(\mathrm{Ded}(\varnothing)) = \mathrm{Ded}(\varnothing),$$

hence $\vdash \sim p \Rightarrow (p \Rightarrow q)$.

This is a mathematical proof of the existence of a proof in the propositional calculus. It is not a proof in the propositional calculus. We shall find other ways of demonstrating the existence of proofs without actually constructing them formally.

Exercises

4.9. Show that $\mathrm{Ded}(A)$ is the smallest subset D of $P(X)$ such that $D \supseteq \mathscr{A} \cup A$ and such that if $p, p \Rightarrow q \in D$, then also $q \in D$.

4.10. Construct a proof in the propositional calculus of $p \Rightarrow r$ from the assumptions $\{p \Rightarrow q, q \Rightarrow r\}$.

We close this chapter with a useful algebraic result.

Theorem 4.11. (The Substitution Theorem). *Let X, Y be any two sets, and let $\varphi: P(X) \to P(Y)$ be a homomorphism of the (free) proposition algebra on X into the (free) proposition algebra on Y. Let $w = w(x_1, \ldots, x_n)$ be any element of $P(X)$ and let A be any subset of $P(X)$. Put $a_i = \varphi(x_i)$.*

(a) *If $A \vdash w$, then $\varphi(A) \vdash w(a_1, \ldots, a_n)$.*

(b) *If $A \vDash w$, then $\varphi(A) \vDash w(a_1, \ldots, a_n)$.*

Proof: (a) Suppose $p_1, \ldots, p_r$ is a proof of w from A. If $p_i \in A$, then trivially $\varphi(p_i) \in \varphi(A)$. Since φ is a homomorphism, it follows that if p_i is an axiom of the propositional calculus on X, then $\varphi(p_i)$ is an axiom of the propositional calculus on Y. For the same reason, if $p_k = (p_j \Rightarrow p_i)$, then $\varphi(p_k) = \varphi(p_j \Rightarrow p_i) = \varphi(p_j) \Rightarrow \varphi(p_i)$. Thus $\varphi(p_1), \ldots, \varphi(p_r)$ is a proof in the propositional calculus on Y of $\varphi(w)$ from $\varphi(A)$. Since $\varphi(w) = w(a_1, \ldots, a_n)$, the result is proved.

(b) Suppose $A \vDash w$. Let $v: P(Y) \to \mathbf{Z}_2$ be a valuation of $P(Y)$ such that $v(\varphi(A)) \subseteq \{1\}$. Then the composite map $v\varphi: P(X) \to \mathbf{Z}_2$ is a valuation of $P(X)$, and $v\varphi(A) \subseteq \{1\}$. Since $A \vDash w$, we have $v\varphi(w) = 1$, i.e. $v(\varphi(w)) = 1$. Thus $\varphi(A) \vDash \varphi(w)$. □

Chapter III

Properties of the Propositional Calculus

§1 Introduction

The properties of the Propositional Calculus that are of interest are those that arise in studying the relation between truth and proof. These properties are important features in the study of any formal system of reasoning, and we begin with some general definitions.

Definition 1.1. A *logic* $\mathscr{L}$ is a system consisting of a set P of elements (called propositions), a set $\mathscr{V}$ of functions (called valuations) from P into some value set W, and, for each subset A of P, a set of finite sequences of elements of P (called proofs from the assumptions A).

For example, the logic called the Propositional Calculus on the set X, and henceforth denoted by Prop(X), consists of the set $P = P(X)$ (the free proposition algebra on X), the set $\mathscr{V}$ of all homomorphisms of $P(X)$ into $\mathbf{Z}_2$, and, for each subset A of $P(X)$, the set of proofs as defined in §4 of Chapter II.

The concepts of semantic implication and syntactic implication in $\mathscr{L}$ are defined in terms of valuation and proof respectively, in some manner analogous to that used for the propositional calculus, and the notations $A \vDash p$, $A \vdash p$ will again be used to denote respectively "p is a consequence of A", "p is a deduction from A". p is a tautology of $\mathscr{L}$ if $\varnothing \vDash p$ and it is a theorem of $\mathscr{L}$ if $\varnothing \vdash p$. The logic $\mathscr{L}$ for which these assertions are made will always be clear from the context.

Definition 1.2. A logic $\mathscr{L}$ is *sound* if $A \vdash p$ implies $A \vDash p$.

Definition 1.3. A logic $\mathscr{L}$ is *consistent* if F is not a theorem.

Definition 1.4. A logic $\mathscr{L}$ is *adequate* if $A \vDash p$ implies $A \vdash p$.

Choosing $A = \varnothing$, we see that a sound logic has the desirable property that theorems are always true, and an adequate logic has the equally desirable property that valid propositions can be proved. While soundness and adequacy each express a connection between truth and proof, consistency is an expression of a purely syntactic property that any logic might be expected to have, namely that one cannot deduce contradictions.

Since the theorems and tautologies of a logic are each of significance, the following decidability properties are also important.

Definition 1.5. A logic $\mathscr{L}$ is *decidable for validity* if there exists an algorithm which determines for every proposition p, in a finite number of steps, whether or not p is valid.

Definition 1.6. A logic $\mathscr{L}$ is *decidable for provability* if there exists an algorithm which determines for every proposition p, in a finite number of steps, whether or not p is a theorem.

§2 Soundness and Adequacy of Prop(X)

Theorem 2.1. (The Soundness Theorem) *Let $A \subseteq P(X)$, $p \in P(X)$. If $A \vdash p$, then $A \vDash p$.*

Proof: Suppose there exists a proof $p_1, \ldots, p_n$ of p from A. We have to show p is a consequence of A.

Let $v: P(X) \to \mathbf{Z}_2$ be a valuation for which $v(A) \subseteq \{1\}$. We shall use induction over the length n of the proof of p from A to show that $v(p) = 1$. Suppose that $n = 1$. Then $p \in A \cup \mathscr{A}$, and since every axiom is a tautology (Exercise 3.8 of Chapter II), we have $v(p) = 1$.

Suppose now $n > 1$, and that $v(q) = 1$ for every q provable from A by a proof of length $< n$. Then $v(p_1) = v(p_2) = \cdots = v(p_{n-1}) = 1$. Either $p_n \in A \cup \mathscr{A}$ and $v(p_n) = 1$, as required, or for some $i, j < n$, we have $p_i = p_j \Rightarrow p_n$. In the latter case, $v(p_j) = v(p_j \Rightarrow p_n) = 1$, and the homomorphism property of v requires $v(p_n) = 1$. □

Corollary 2.2. (The Consistency Theorem) *F is not a theorem of* Prop(X).

Proof: If $\vdash F$, then $\vDash F$ by the Soundness Theorem. Since axioms are tautologies, $v(F) = 1$ for every valuation v, contradicting the definition of valuation. This implies that there are no valuations. But $P(X)$ is free and every map of X into $\mathbf{Z}_2$ can be extended to a valuation. □

Exercise 2.3. Show that Con(A) is closed with respect to modus ponens (i.e., if $p, p \Rightarrow q \in$ Con(A), then $q \in$ Con(A)). Use Exercise 4.9 of Chapter II to prove that Con(A) $\supseteq$ Ded(A). This is another way of stating the Soundness Theorem.

The proof of adequacy for Prop(X) is more difficult, and we first prove a preparatory result of independent interest.

Theorem 2.4. (The Deduction Theorem) *Let $A \subseteq P(X)$, and let p, $q \in P(X)$. Then $A \vdash p \Rightarrow q$ if and only if $A \cup \{p\} \vdash q$.*

Proof: (a) Suppose $A \vdash p \Rightarrow q$. Let $p_1, \ldots, p_n$ be a proof of $p_n = p \Rightarrow q$ from A. Then $p_1, \ldots, p_n, p, q$ is a proof of q from $A \cup \{p\}$.

(b) Suppose $A \cup \{p\} \vdash q$. Then we have a proof $p_1, \ldots, p_n$ of q from $A \cup \{p\}$. We shall use induction over the length n of the proof.

If $n = 1$, then $q \in \mathscr{A} \cup A \cup \{p\}$. If $q \in \mathscr{A} \cup A$, then q, $q \Rightarrow (p \Rightarrow q)$, $p \Rightarrow q$ is a proof of $p \Rightarrow q$ from A. If $q = p$, then $\vdash p \Rightarrow p$ (Example 3.4 of Chapter II), and so $A \vdash p \Rightarrow q$.

Suppose now $n > 1$. By induction, $A \vdash p \Rightarrow p_i$ for $i = 1, 2, \ldots, n - 1$, and we may suppose $q \notin \mathscr{A} \cup A \cup \{p\}$. For some $i, j < n$, we have $p_i = p_j \Rightarrow q$. Thus $A \vdash p \Rightarrow p_j$, $A \vdash p \Rightarrow (p_j \Rightarrow q)$, and there exists a proof $q_1, \ldots, q_k$, q_{k+1} from A with

$$\begin{aligned} q_k &= p \Rightarrow p_j, \\ q_{k+1} &= p \Rightarrow (p_j \Rightarrow q). \end{aligned}$$

We put

$$\begin{aligned} q_{k+2} &= (p \Rightarrow (p_j \Rightarrow q)) \Rightarrow ((p \Rightarrow p_j) \Rightarrow (p \Rightarrow q)), && (\mathscr{A}_2) \\ q_{k+3} &= (p \Rightarrow p_j) \Rightarrow (p \Rightarrow q), && (q_{k+2} = q_{k+1} \Rightarrow q_{k+3}) \\ q_{k+4} &= p \Rightarrow q. && (q_{k+3} = q_k \Rightarrow q_{k+4}) \end{aligned}$$

Then $q_1, \ldots, q_{k+4}$ is a proof of $p \Rightarrow q$ from A. □

The Deduction Theorem is useful in establishing a result of the form $A \vdash p \Rightarrow q$, because it is usually much easier to show $A \cup \{p\} \vdash q$. Even if a proof in Prop(X) of $p \Rightarrow q$ from A is required, the method used in proving the Deduction Theorem can be applied to convert a proof of q from $A \cup \{p\}$ into a proof of $p \Rightarrow q$ from A.

Example 2.5. We show $\{p \Rightarrow q, q \Rightarrow r\} \vdash p \Rightarrow r$. First we show $\{p, p \Rightarrow q, q \Rightarrow r\} \vdash r$, and a proof of this is p, $p \Rightarrow q$, q, $q \Rightarrow r$, r. It follows from the Deduction Theorem that $\{p \Rightarrow q, q \Rightarrow r\} \vdash p \Rightarrow r$.

We now convert the proof of r from $\{p, p \Rightarrow q, q \Rightarrow r\}$ into a proof of $p \Rightarrow r$ from $\{p \Rightarrow q, q \Rightarrow r\}$. We shall write the steps of the original proof in a column on the left. Alongside each, we then write a comment on the nature of the step, and then the corresponding steps of the new proof.

p	Proposition to be deleted from the assumptions	$p \Rightarrow ((p \Rightarrow p) \Rightarrow p)$, $(p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p))$, $(p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)$, $p \Rightarrow (p \Rightarrow p)$, $p \Rightarrow p$.
$p \Rightarrow q$	Retained assumption	$p \Rightarrow q$, $(p \Rightarrow q) \Rightarrow (p \Rightarrow (p \Rightarrow q))$, $p \Rightarrow (p \Rightarrow q)$.
q	Modus ponens	$(p \Rightarrow (p \Rightarrow q)) \Rightarrow ((p \Rightarrow p) \Rightarrow (p \Rightarrow q))$, $(p \Rightarrow p) \Rightarrow (p \Rightarrow q)$, $p \Rightarrow q$.
$q \Rightarrow r$	Retained assumption	$q \Rightarrow r$, $(q \Rightarrow r) \Rightarrow (p \Rightarrow (q \Rightarrow r))$, $p \Rightarrow (q \Rightarrow r)$.
r	Modus ponens	$(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r))$, $(p \Rightarrow q) \Rightarrow (p \Rightarrow r)$, $p \Rightarrow r$.

Of course, the proof we have constructed can be abbreviated, because the first 11 steps serve only to prove the retained assumption $p \Rightarrow q$.

Exercises

2.6. Show that $p \Rightarrow r \in \mathrm{Ded}\{p \Rightarrow q,\ p \Rightarrow (q \Rightarrow r)\}$. Hence show that if $p \Rightarrow q, p \Rightarrow (q \Rightarrow r) \in \mathrm{Ded}(A)$, then $p \Rightarrow r \in \mathrm{Ded}(A)$, and so prove the Deduction Theorem without giving an explicit construction for a proof in Prop(X).

2.7. Show that $\vdash p \Rightarrow \sim \sim p$ and construct a proof of $p \Rightarrow \sim \sim p$ in Prop(X). (Hint: show $\{p, \sim p\} \vdash F$ and use the Deduction Theorem twice.)

2.8. Show that the following are theorems of Prop(X),

(a) $p \Rightarrow p \vee q$, (b) $q \Rightarrow p \vee q$,
(c) $(p \vee q) \Rightarrow (q \vee p)$, (d) $p \wedge q \Rightarrow p$,
(e) $p \wedge q \Rightarrow q$, (f) $(p \wedge q) \Rightarrow (q \wedge p)$.

Definition 2.9. Let $A \subseteq P(X)$. We say that A is *consistent* if $F \notin \mathrm{Ded}(A)$. A is called a *maximal consistent* subset if A is consistent and if every subset $T \subseteq P(X)$ which properly contains A is inconsistent.

Lemma 2.10. *The subset $A \subseteq P(X)$ is maximal consistent if and only if*
(i) *$F \notin A$, and*
(ii) *$A = \mathrm{Ded}(A)$, and*
(iii) *for all $p \in P(X)$, either $p \in A$ or $\sim p \in A$.*

Proof: (a) Let A be maximal consistent. Since A is consistent, $F \notin \mathrm{Ded}(A)$ and therefore $F \notin A$. Since $\mathrm{Ded}(\mathrm{Ded}(A)) = \mathrm{Ded}(A)$, $\mathrm{Ded}(A)$ is consistent. As $A \subseteq \mathrm{Ded}(A)$, $A = \mathrm{Ded}(A)$ by the maximal consistency of A. Finally, suppose $p \notin A$. Then $F \in \mathrm{Ded}(A \cup \{p\})$, i.e. $A \cup \{p\} \vdash F$. By the Deduction Theorem, $A \vdash p \Rightarrow F$, i.e., $\sim p \in \mathrm{Ded}(A)$.

(b) Suppose A has the properties (i), (ii), (iii). Then $F \notin \mathrm{Ded}(A)$. If T properly contains A, then there exists $p \in T$ such that $p \notin A$. By (iii), $\sim p \in A$, hence $p, \sim p \in T$, and $p, \sim p, F$ is a proof of F from T. Thus A is maximal consistent. □

Lemma 2.11. *Let A be a consistent subset of $P(X)$. Then A is contained in a maximal consistent subset.*

Proof: Let $\Sigma = \{T \subseteq P(X) \mid T \supseteq A, F \notin \mathrm{Ded}(T)\}$. Since $A \in \Sigma$, $\Sigma \neq \varnothing$. Suppose $\{T_\alpha\}$ is a totally ordered family of members of Σ, and put $T = \bigcup_\alpha T_\alpha$. Clearly $T \subseteq P(X)$, $T \supseteq A$. If F is provable from T, F is provable from a finite subset of T, and this subset is contained in some T_α, contrary to $T_\alpha \in \Sigma$. Hence $F \notin \mathrm{Ded}(T)$, and Σ is an inductively ordered set. By Zorn's Lemma, Σ has a maximal member say M. This M is the required maximal consistent subset. □

The next result is the key to the Adequacy Theorem.

Theorem 2.12. (The Satisfiability Theorem) *Let A be a consistent subset of $P(X)$. Then there exists a valuation $v: P(X) \to \mathbb{Z}_2$, such that $v(A) \subseteq \{\mathfrak{t}\}$.*

Proof: Let M be a maximal consistent subset containing A. For $p \in P(X)$, put $v(p) = 1$ if $p \in M$ and $v(p) = 0$ if $p \notin M$. We now prove v is a valuation.

Certainly $v(F) = 0$, because $F \notin M$. It remains to show $v(p \Rightarrow q) = v(p) \Rightarrow v(q)$. If $q \in M$, then $p \Rightarrow q \in M$ because $\{q\} \vdash p \Rightarrow q$, and $v(p \Rightarrow q) = 1 = v(p) \Rightarrow v(q)$. If $p \notin M$, then $p \Rightarrow q \in M$ because $\{\sim p\} \vdash p \Rightarrow q$, and $v(p \Rightarrow q) = 1 = v(p) \Rightarrow v(q)$. If $p \in M$ and $q \notin M$, then $p \Rightarrow q \notin M$, and $v(p \Rightarrow q) = 0 = v(p) \Rightarrow v(q)$. □

Theorem 2.13. (The Adequacy Theorem) *Let* $A \subseteq P(X)$, $p \in P(X)$. *If* $A \vDash p$ *in* Prop(X), *then* $A \vdash p$ *in* Prop(X).

Proof: Suppose $A \vDash p$, so that $v(A) \subseteq \{1\}$ implies $v(p) = 1$ for every valuation v. If $A \cup \{\sim p\}$ is consistent, it follows from the Satisfiability Theorem that there is a valuation v such that $v(A \cup \{\sim p\}) \subseteq \{1\}$, which is not possible. Hence $F \in \text{Ded}(A \cup \{\sim p\})$, i.e., $A \cup \{\sim p\} \vdash F$. By the Deduction Theorem, $A \vdash \sim p \Rightarrow F$. Since $\vdash \sim \sim p \Rightarrow p$, we have $A \vdash p$. □

Exercise 2.14. Show that if $A \vDash p$, then $A_0 \vDash p$ for some finite subset A_0 of A. (This result is known as the Compactness Theorem.)

§3 Truth Functions and Decidability for Prop(X)

Each valuation v of $P(X)$ determines a natural equivalence relation r_v on $P(X)$ given by pr_vq if $v(p) = v(q)$, and which is in fact a congruence relation on $P(X)$. That is, each r_v satisfies the condition that if pr_vp_1 and qr_vq_1, then $(p \Rightarrow q)r_v(p_1 \Rightarrow q_1)$. The intersection of the relations r_v for all valuations v of $P(X)$ is therefore a congruence relation on $P(X)$, which we call semantic equivalence and denote by $\vDash\!\dashv$. Since $p \vDash\!\dashv q$ if and only if $v(p) = v(q)$ for every valuation v of $P(X)$, we see that $p \vDash\!\dashv q$ if and only if $\{p\} \vDash q$ and $\{q\} \vDash p$.

Definition 3.1. The set of congruence classes of $P(X)$ with respect to $\vDash\!\dashv$ is an $\{F, \Rightarrow\}$—algebra called the *Lindenbaum algebra* on X and denoted by $L(X)$.

Let $X_n = \{x_1, \ldots, x_n\}$. Clearly $L(X_n)$ is a homomorphic image of $P(X_n)$. If $w = w(x_1, \ldots, x_n) \in P(X_n)$ is any word in $x_1, \ldots, x_n$, then its image in $L(X_n)$ is the congruence class $\bar{w} = \bar{w}(x_1, \ldots, x_n)$ say, of all words congruent to w under the relation $\vDash\!\dashv$. Our aim is to show that $\bar{w}$ can be regarded as a function $\bar{w}: Z_2^n \to Z_2$.

For any $\bar{w}(x_1,\ldots,x_n) \in L(X_n)$, choose a representative $w(x_1,\ldots,x_n) \in P(X_n)$. If $(z_1, \ldots, z_n) \in Z_2^n$, then there is a unique valuation $v: P(X_n) \to Z_2$ such that $v(x_i) = z_i$ for $i = 1, \ldots, n$. We define $\bar{w}(z_1, \ldots, z_n) = v(w(x_1, \ldots, x_n))$, observing that this definition is independent of the choice of representative w of $\bar{w}$, because if w_1 is another representative, then $w \vDash\!\dashv w_1$ and $v(w) = v(w_1)$. In this way we associate with each element $\bar{w}$ of $L(X_n)$ a function $Z_2^n \to Z_2$, but, before we identify $\bar{w}$ with this function, we must show that if $\bar{w}$ and $\bar{w}_1$ have the same associated function, then $\bar{w} = \bar{w}_1$.

Suppose that $\bar{w}$ and $\bar{w}_1$ have the same associated function, so that $\bar{w}(z_1, \ldots, z_n) = \bar{w}_1(z_1, \ldots, z_n)$ for all $(z_1, \ldots, z_n) \in \mathbf{Z}_2^n$. Let w, w_1 be representatives of $\bar{w}, \bar{w}_1$ respectively. Then $\bar{w}(z_1, \ldots, z_n) = v(w)$, where v is the valuation for which $v(x_i) = z_i$ $(i = 1, \ldots, n)$, and we have $v(w) = v(w_1)$. The last equation holds for every valuation v, hence $w \models\!\dashv w_1$ and $\bar{w} = \bar{w}_1$. We may therefore identify the elements of $L(X_n)$ with their associated functions.

Definition 3.2. A function $f: \mathbf{Z}_2^n \to \mathbf{Z}_2$ is called a truth function.

Theorem 3.3. *$L(X_n)$ is the set of all truth functions $f: \mathbf{Z}_2^n \to \mathbf{Z}_2$.*

Proof: The constant functions $0, 1 \in L(X_n)$ since $0 = \bar{F}$ and $1 = \overline{(F \Rightarrow F)}$. Thus the result holds for $n = 0$.

If f, g are truth functions $\mathbf{Z}_2^n \to \mathbf{Z}_2$, we define the truth function $f \Rightarrow g$ by $(f \Rightarrow g)(z_1, \ldots, z_n) = f(z_1, \ldots, z_n) \Rightarrow g(z_1, \ldots, z_n)$. For convenience of notation, we denote the ith coordinate function by u_i. We have $u_i = \bar{x}_i \in L(X_n)$.

We now suppose $n > 0$, and shall use induction over n to complete the proof. Let $f = f(u_1, \ldots, u_n)$ be a truth function of n variables. Put

$$g(u_1, \ldots, u_{n-1}) = f(u_1, \ldots, u_{n-1}, 0), h(u_1, \ldots, u_{n-1}) = f(u_1, \ldots, u_{n-1}, 1).$$

Then $g, h \in L(X_{n-1}) \subseteq L(X_n)$. The function $k: \mathbf{Z}_2^n \to \mathbf{Z}_2$, defined by

$$k(u_1, \ldots, u_n) = (\sim u_n \Rightarrow g(u_1, \ldots, u_{n-1})) \wedge (u_n \Rightarrow h(u_1, \ldots, u_{n-1}))$$

is in $L(X_n)$, and

$$\begin{aligned} k(u_1, \ldots, u_{n-1}, 0) &= (1 \Rightarrow g(u_1, \ldots, u_{n-1})) \wedge (0 \Rightarrow h(u_1, \ldots, u_{n-1})) \\ &= g(u_1, \ldots, u_{n-1}) \wedge 1 \\ &= g(u_1, \ldots, u_{n-1}) \\ &= f(u_1, \ldots, u_{n-1}, 0). \end{aligned}$$

Similarly, one obtains $k(u_1, \ldots, u_{n-1}, 1) = f(u_1, \ldots, u_{n-1}, 1)$. Thus $k = f$ and $f \in L(X_n)$. □

We now apply truth functions to settle the question of decidability for Prop(X).

Lemma 3.4. *Let $w = w(x_1, \ldots, x_n) \in P(X)$. Then $\models w$ if and only if its associated truth function $\bar{w}: \mathbf{Z}_2^n \to \mathbf{Z}_2$ is the constant 1.*

Proof: Suppose $\bar{w} = 1$. Let $v: P(X) \to \mathbf{Z}_2$ be any valuation of $P(X)$. Put $a_i = v(x_i)$. Then the restriction of v to $P(X_n)$ is a valuation of $P(X_n)$, and $v(w) = \bar{w}(a_1, \ldots, a_n) = 1$. Thus $v(w) = 1$ for every valuation v of $P(X)$, i.e., $\models w$.

Suppose conversely that w is valid. Let $(a_1, \ldots, a_n) \in \mathbf{Z}_2^n$. There exists a valuation v of $P(X)$ with $v(x_i) = a_i$. (We may assign arbitrarily values for elements of $X - \{x_1, \ldots, x_n\}$.) Then the restriction of v to $P(X_n)$ is a valuation of $P(X_n)$, and $\bar{w}(a_1, \ldots, a_n) = v(w) = 1$. Thus $\bar{w} = 1$. □

Theorem 3.5. Prop(X) *is decidable for validity.*

Proof: We give an algorithm for deciding if $w \in P(X)$ is valid. The element w is a word $w(x_1, \ldots, x_n)$ in some finite set $x_1, \ldots, x_n$ of variables. Let $\bar{w} = w(u_1, \ldots, u_n)$ be the associated truth function. For each $(a_1, \ldots, a_n) \in \mathbf{Z}_2^n$, we calculate $\bar{w}(a_1, \ldots, a_n)$. By Lemma 3.4, w is valid if and only if all these values are 1. □

Corollary 3.6. Prop(X) *is decidable for provability.*

Proof. An element $p \in P(X)$ is a theorem if and only if it is valid. □

Exercises

3.7. Show that every truth function $\mathbf{Z}_2^n \to \mathbf{Z}_2$ can be expressed in terms of the coordinate functions and the one operation $|$ defined by $\bar{w}_1|\bar{w}_2 = \sim(\bar{w}_1 \wedge \bar{w}_2)$.

3.8. A truth function $f(u_1, \ldots, u_n)$ is said to be in disjunctive normal form if it is expressed in one of the forms $f = 0, f = 1$, or $f = v_1 \vee v_2 \vee \cdots \vee v_k$ for $0 < k < 2^n$, where each $v_j = u_{1j} \wedge u_{2j} \wedge \cdots \wedge u_{n_j j}$, and $u_{ij} = u_r$ or $\sim u_r$ for some r.

Show that every truth function is expressible in disjunctive normal form, and specify a procedure for associating with each truth function $\mathbf{Z}_2^n \to \mathbf{Z}_2$ a unique disjunctive normal form.

3.9. (a) Let $p \in P(X)$. Find a $p' \in P(X)$, expressible in a form involving no operations other than $\sim$, $\wedge$ and $\vee$, such that $\models p \Leftrightarrow p'$.

(b) Let $p, q \in P(X)$. Find truth functions for $\sim(p \vee q) \Leftrightarrow (\sim p \wedge \sim q)$ and $\sim(p \wedge q) \Leftrightarrow (\sim p \vee \sim q)$.

(c) p and p' are related as in (a). Let p^* be the statement obtained from p' by replacing each $\vee$ by $\wedge$, each $\wedge$ by $\vee$, and each $x \in X$ by $\sim x$. Prove that $\models \sim p \Leftrightarrow p^*$.

3.10. A truth function $f(u_1, \ldots, u_n)$ is said to be in conjunctive normal form if it expressed in one of the forms $f = 0, f = 1$, or $f = v_1 \wedge v_2 \wedge \cdots \wedge v_k$ for $0 < k < 2^n$, where each $v_j = u_{1j} \vee u_{2j} \vee \cdots \vee u_{n_j j}$, and $u_{ij} = u_r$ or $\sim u_r$ for some r. Use Exercises 3.8 and 3.9 to specify a procedure for associating with each truth function $\mathbf{Z}_2^n \to \mathbf{Z}_2$ a unique conjunctive normal form.

3.11. Let p, p' and q, q' be related as in Exercise 3.9(a). Let p^d, q^d be the statements obtained from p', q' by replacing each $\vee$ by $\wedge$ and each $\wedge$ by $\vee$. Show that $\models p$ if and only if $\models \sim p^d$. Deduce that if $\vdash p \Rightarrow q$, then $\vdash q^d \Rightarrow p^d$. (This result expresses a duality principle for Prop(X).)

3.12. Write a FORTRAN program to decide if $w(x_1, x_2, x_3) \in P(X_3)$ is valid.

3.13. Show that Prop(X) is decidable for $\{p_1, \ldots, p_n\} \models q$, where $p_1, \ldots, p_n, q \in P(X)$.

3.14. Construct a propositional calculus Prop$_1(X)$ with $P_1(X)$ the free $\{\Rightarrow, \sim\}$-algebra. Show that there is a $\{\Rightarrow, \sim\}$-homomorphism φ: $P_1(X) \to P(X)$ which is the identity on X. Is φ a monomorphism? Is φ an epimorphism? Does there exist a $\{\Rightarrow, \sim\}$-homomorphism $\psi: P(X) \rightarrow P_1(X)$ which is the identity on X? (Hint: Consider the images of F and of $F \Rightarrow F$ $(= \sim F)$.)

Show that there exists a $\{\Rightarrow, F\}$-homomorphism $\theta: P(X) \rightarrow P_1(X)$ which is the identity on X, taking as element F of $P_1(X)$ the element $\sim(x_1 \Rightarrow x_1)$. Show that $w \in P_1(X)$ is valid if and only if $\varphi(w)$ is valid. Show that $p \in P(X)$ is valid if and only if $\theta(p)$ is valid. Establish the Consistency, Adequacy and Decidability theorems for $\text{Prop}_1(X)$.

3.15. Using the method of 3.14 investigate the following propositional calculi:

(a) $\text{Prop}_2(X)$ with $P_2(X)$ free of type $\{\sim, \vee\}$,

(b) $\text{Prop}'_2(X)$, with $P'_2(X)$ relatively free of type $\{\sim, \vee\}$, with the identical relation $p \vee q = q \vee p$,

(c) $\text{Prop}_3(X)$ with $P_3(X)$ free of type $\{|\}$ (see 3.7),

(d) $\text{Prop}'_3(X)$ with $P'_3(X)$ relatively free of type $\{|\}$, with the identical relation $p|q = q|p$.

Chapter IV

Predicate Calculus

§1 Algebras of Predicates

The initial step in our development of the Propositional Calculus was the construction of proposition algebras, which formalise the way in which a given collection of "primitive" statements is enlarged by combining statements. The Propositional Calculus does not analyse the original primitive statements. Our aim now is to construct a more complicated model of mathematical reasoning, which incorporates more of the ordinary features of this reasoning.

Mathematics is usually about something, that is, there is usually some set $\mathscr{U}$ of objects under discussion and investigation. A typical statement in such a discussion would be "u has the property p", where $u \in \mathscr{U}$ and p is some property relevant to elements of $\mathscr{U}$. A convenient notation for this statement is $p(u)$. Such a statement depends on the element u, and may be thought of as a function of u. The phrase "has the property p" is known as a predicate, and p (as used in the notation $p(u)$) is known as a predicate symbol. More generally, if r is an n-ary relation on $\mathscr{U}$, the statement "$(u_1, \ldots, u_n)$ is in the relation r" is denoted by $r(u_1, \ldots, u_n)$, and r is called an n-ary predicate. A 0-ary predicate is a statement which does not depend on any elements of $\mathscr{U}$, and so corresponds to an unanalysed statement.

If p, q are properties, then $p(u) \wedge q(u)$ is true for just those elements u with both properties. Denoting by P the subset of $\mathscr{U}$ consisting of those elements with property p, and by Q the subset of $\mathscr{U}$ of elements with property q, we see that $P \cap Q$ is the subset of those elements u for which $p(u) \wedge q(u)$ is true. Similarly, $P \cup Q$ is the subset of elements u for which $p(u) \vee q(u)$ is true, while the set of elements u satisfying $\sim p(u)$ is the complement of P in $\mathscr{U}$.

Another common form of statement in mathematical discussion is "For all $u \in \mathscr{U}$, $p(u)$". If $\mathscr{U}$ were a finite set, say $\mathscr{U} = \{u_1, \ldots, u_n\}$, then this could be expressed as $p(u_1) \wedge p(u_2) \wedge \cdots \wedge p(u_n)$, but it is not possible to do this if $\mathscr{U}$ is an infinite set. We thus introduce the notation $(\forall u)p(u)$ for the above statement. $(\forall u)$ is called the universal quantifier. Note that the u in $(\forall u)$ is only a dummy—$(\forall u)p(u)$ is in no way dependent on u, and is the same statement about $\mathscr{U}$ as $(\forall v)p(v)$. We do not need additional notations to deal with a limited use of "for all" as in statements such as "For all u such that $p(u)$, we have $q(u)$". This can be expressed as $(\forall u)(p(u) \Rightarrow q(u))$.

Statements of the form "There exists $u \in \mathscr{U}$ with the property p" are also common in mathematics. We write this statement as $(\exists u)p(u)$. The existential quantifier $(\exists u)$ is, however, related to the universal quantifier $(\forall u)$, as follows. When we say "There does not exist u with property p", we are in fact asserting $(\forall u)(\sim p(u))$. Thus $(\exists u)p(u)$ has the same meaning[1] as $\sim((\forall u)(\sim p(u))$, and we have no need to include the existential quantifier in the construction of our model. We shall define $(\exists u)$ to mean $\sim(\forall u)\sim$.

We now set up an appropriate analogue of a proposition algebra. Proposition algebras are built upon underlying sets of propositional variables. We begin here with an infinite set V whose elements will be called individual variables, and with a set $\mathscr{R}$ (whose elements will be called relation or predicate symbols) together with an arity function $\text{ar}: \mathscr{R} \to \mathbf{N}$. The individual variables may be thought of as names to be given to mathematical objects, and the relation symbols as names to be given to relations between these objects. The set of generators we shall use to construct our set P of propositions must clearly contain each element $r(x_1, \ldots, x_n)$ for each $r \in \mathscr{R}$ and $(x_1, \ldots, x_n) \in V^n$, where $n = \text{ar}(r)$. It is also clear that P must be an $\{F, \Rightarrow\}$—algebra, and that for each $x \in V$, we shall need a function $(\forall x): P \to P$.

Let $\tilde{P}(V, \mathscr{R})$ be the free algebra on the set $\{(r, x_1, \ldots, x_n) | r \in \mathscr{R}, x_i \in V, n = \text{ar}(r)\}$ of free generators, of type $\{F, \Rightarrow, (\forall x) | x \in V\}$, where F is a 0-ary operation, $\Rightarrow$ binary, and each $(\forall x)$ unary. We call $\tilde{P} = \tilde{P}(V, \mathscr{R})$ the full first order algebra on $(V, \mathscr{R})$. We use the more usual notation $r(x_1, \ldots, x_n)$ for the generator $(r, x_1, \ldots, x_n)$, and we put $\mathscr{R}_n = \{r \in \mathscr{R} | \text{ar}(r) = n\}$.

We could use this algebra $\tilde{P}$ as our algebra of propositions, but it is more convenient to use a certain factor algebra. If $w \in \tilde{P}$, then w is a word in the free generators of $\tilde{P}$, each of which has the form $r(x_1, \ldots, x_n)$. If $x_1, \ldots, x_m$ are the distinct individual variables occurring in w, then we can think of w as a function $w(x_1, \ldots, x_m)$ of these variables. Now we regard $(\forall x_1)w(x_1, \ldots, x_m)$ as being essentially the same as $(\forall y)w(y, x_2, \ldots, x_m)$, provided only that $y \notin \{x_2, \ldots, x_m\}$. The reason for this has been pointed out before, and is that the x_1 in $(\forall x_1)w(x_1, \ldots, x_m)$ is a dummy, used as an aid in describing the construction of the statement. It serves the same purpose as the variable t does in the definition of the gamma function as $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}\,dt$.

We shall construct a factor algebra of $\tilde{P}$, in which these elements, considered above as being essentially the same, will be identified. Further identifications are possible. The question of which identifications are made is purely one of convenience. The congruence relation on $\tilde{P}$ which we use needs some care in its construction, and we begin by defining two functions on $\tilde{P}$.

Definition 1.1. Let $w \in \tilde{P}$. The set of *variables involved in* w, denoted by $V(w)$, is defined by

$$V(w) = \cap \{U | U \subseteq V, w \in \tilde{P}(U, \mathscr{R})\}.$$

[1] This is very different to the concepts of existence used in other contexts such as "Do flying saucers exist?" or "Does God exist?" or "Do electrons exist?".

Exercise 1.2. Show that

(i) $V(F) = \varnothing$.

(ii) If $r \in \mathscr{R}$, $\operatorname{ar}(r) = n$, and $x_1, \ldots, x_n \in V$, then $V(r(x_1, \ldots, x_n)) = \{x_1, \ldots, x_n\}$.

(iii) If $w_1, w_2 \in \tilde{P}$, then $V(w_1 \Rightarrow w_2) = V(w_1) \cup V(w_2)$.

(iv) If $x \in V$ and $w \in \tilde{P}$, then $V((\forall x)w) = \{x\} \cup V(w)$.

Show further that (i)–(iv) may be taken as the definition of the function $V(w)$.

Definition 1.3. Let $w \in \tilde{P}$. The *depth of quantification* of w, denoted by $d(w)$, is defined by

(i) $d(F) = 0$, $d(r(x_1, \ldots, x_n)) = 0$ for every free generator of $\tilde{P}$.

(ii) $d(w_1 \Rightarrow w_2) = \max(d(w_1), d(w_2))$.

(iii) $d((\forall x)w) = 1 + d(w)\ (x \in V)$.

Our desired congruence relation on $\tilde{P}$ may now be defined.

Definition 1.4. Let $w_1, w_2 \in \tilde{P}$. We define $w_1 \approx w_2$ if

(a) $d(w_1) = d(w_2) = 0$ and $w_1 = w_2$, or

(b) $d(w_1) = d(w_2) > 0, w_1 = a_1 \Rightarrow b_1, w_2 = a_2 \Rightarrow b_2, a_1 \approx a_2$ and $b_1 \approx b_2$, or

(c) $w_1 = (\forall x)a$, $w_2 = (\forall y)b$ and either

(i) $x = y$ and $a \approx b$, or

(ii) there exists $c = c(x)$ such that $c(x) \approx a$, $c(y) \approx b$ and $y \notin V(c)$.

We remark that in part (c) (ii), the notation $c = c(x)$ indicates the way the element concerned is a function of x, and ignores its possible dependence on other variables. We use it so we can represent the effect of substituting y for x throughout. It is therefore unnecessary for us to impose the condition $x \notin V(c(y))$. The notation does not imply $V(c(x)) = \{x\}$, hence we must impose the condition $y \notin V(c(x))$. Thus the condition (c) (ii) is symmetric, and $\approx$ is trivially reflexive. The proof that it is transitive is left as an exercise.

Exercise 1.5.

(i) Given that $z \notin V(w_1) \cup V(w_2)$, show by induction over $d(w_1)$ that the element $c = c(x)$ in (c) (ii) can always be chosen such that $z \notin V(c)$.

(ii) If $u(x) \approx v(x)$ and $y \notin V(u(x)) \cup V(v(x))$, show by induction over $d(u(x))$ that $u(y) \approx v(y)$.

(iii) Prove that $\approx$ is transitive.

Since the relation $\approx$ is an equivalence which is clearly compatible with the operations of the algebra, it is a congruence relation on $\tilde{P}(V, \mathscr{R})$.

Definition 1.6. The *(reduced) first-order algebra* $P(V, \mathscr{R})$ on $(V, \mathscr{R})$ is the factor algebra of $\tilde{P}(V, \mathscr{R})$ by the congruence relation $\approx$.

The elements of $P = P(V, \mathscr{R})$ are the congruence classes. If $w \in \tilde{P}$ and

$[w]$ is the congruence class of w, then

$$(\forall x)[w] = [(\forall x)w],$$

and

$$[w_1] \Rightarrow [w_2] = [w_1 \Rightarrow w_2],$$

Definition 1.7. Let $w \in P$. We define the set $\operatorname{var}(w)$ of *(free) variables of* w by putting $\operatorname{var}(w) = \operatorname{var}(\tilde{w})$, where $\tilde{w} \in \tilde{P}$ is some representative of the congruence class w, and where $\operatorname{var}(\tilde{w})$ is defined inductively by

(i) $\operatorname{var}(F) = \varnothing$,
(ii) $\operatorname{var}(r(x_1, \ldots, x_n)) = \{x_1, \ldots, x_n\}$ for $r \in \mathscr{R}$, $x_1, \ldots, x_n \in V$,
(iii) $\operatorname{var}(\tilde{w}_1 \Rightarrow \tilde{w}_2) = \operatorname{var}(\tilde{w}_1) \cup \operatorname{var}(\tilde{w}_2)$,
(iv) $\operatorname{var}((\forall x)\tilde{w}) = \operatorname{var}(\tilde{w}) - \{x\}$.

Definition 1.8. Let $A \subseteq P$. Put

$$\operatorname{var}(A) = \bigcup_{p \in A} \operatorname{var}(p).$$

Exercises

1.9. Show that if $\tilde{w}_1 \approx \tilde{w}_2$, then $\operatorname{var}(\tilde{w}_1) = \operatorname{var}(\tilde{w}_2)$, and conclude that $\operatorname{var}(w)$ is defined for $w \in P$.

1.10. Show that for any $w \in P$, there is a representative $\tilde{w}$ of w such that no variable $x \in V$ appears in $\tilde{w}$ more than once in a quantifier $(\forall x)$, and no $x \in \operatorname{var}(w)$ appears at all in a quantifier (i.e., $\tilde{w}$ has no repeated dummy variables, and no free variables also appear as dummies).

We assume henceforth that any $w \in P$ is represented by a $\tilde{w} \in \tilde{P}$ having the form described in Exercise 1.10. We shall also usually abuse notation and not distinguish between $p \in \tilde{P}$ and $[p] \in P$.

§2 Interpretations

We want to think of the elements of V as names of objects, and the elements of $\mathscr{R}$ as relations among those objects. If we take a non-empty set U, and a function $\varphi: V \to U$, then we can think of $x \in V$ as a name for the element $\varphi(x) \in U$. Of course, not every element $u \in U$ need have a name, while some elements u may well have more than one name. Next we take a function ψ, from $\mathscr{R}$ into the set of all relations on U, such that if $r \in \mathscr{R}_n$, then $\psi(r)$ is an n-ary relation. It will be convenient to write simply φx for $\varphi(x)$, and ψr for $\psi(r)$. As for valuations, these again should be functions $v: P \to \mathbf{Z}_2$ which will correspond to our intuitive notion of truth. Since our interpretation of the element $r(x_1, \ldots, x_n) \in P$ in terms of U, φ, ψ must obviously be the statement that $(\varphi x_1, \ldots, \varphi x_n) \in \psi r$, we shall require of v that

(a) if $r \in \mathscr{R}_n$ and $x_1, \ldots, x_n \in V$, then $v(r(x_1, \ldots, x_n) = 1$ if $(\varphi x_1, \ldots, \varphi x_n) \in \psi r$, and is 0 otherwise, while we still require that

(b) v is a homomorphism of $\{F, \Rightarrow\}$-algebras.

It remains for us to define truth for a proposition of the form $(\forall x)p(x)$ in terms of our understanding of it for $p(x)$, and so we use an induction over the depth of quantification. Let $P_k(V, \mathscr{R})$ be the set of all elements p of $P(V, \mathscr{R})$ with $d(p) \leqslant k$. If we take some new variable t, then intuitively, we consider $(\forall x)p(x)$ $(=(\forall t)p(t))$ to be true if $p(t)$ is true no matter how we choose to interpret t. This leads to a further requirement for v, namely:

(c_k) Suppose $p = (\forall x)q(x)$ has depth k. Put $V' = V \cup \{t\}$ where $t \notin V$. If for every extension $\varphi': V' \to U$ of φ and for every $v'_{k-1}: P_{k-1}(V', R) \to \mathbf{Z}_2$, such that $(\varphi', \psi, v'_{k-1})$ satisfy (a), (b) and (c_i) for all $i < k$, we have $v'_{k-1}(q(t)) = 1$, then $v(p) = 1$, otherwise $v(p) = 0$.

Exercise 2.1. Given U, φ, ψ, prove that there is one and only one function $v: P \to \mathbf{Z}_2$ satisfying (a), (b) and (c_i) for all i.

Briefly, the above exposition of the components of an interpretation of $P(V, \mathscr{R})$ can be expressed as follows.

Definition 2.2. An *interpretation* of $P = P(V, \mathscr{R})$ in the domain U is a quadruple (U, φ, ψ, v) satisfying the conditions (a), (b) and (c_k) for all k.

As before, we write $A \vDash p$ if $A \subseteq P$, $p \in P$ and $v(p) = 1$ for every interpretation of P for which $v(A) \subseteq \{1\}$. We denote by Con(A) the set of all p such that $A \vDash p$. We write $\vDash p$ for $\varnothing \vDash p$, and any p for which $\vDash p$, is called valid or a tautology.

Exercises

2.3. Let $w(u_1, \ldots, u_n)$ be any tautology of Prop($\{u_1, \ldots, u_n\}$). Let $p_1, \ldots, p_n \in P(V, \mathscr{R})$. Prove that $\vDash w(p_1, \ldots, p_n)$.

2.4. $A \subseteq P(V, \mathscr{R})$ and $p(x) \in A$ for all $x \in V$. Does it follow that $A \vDash (\forall x)p(x)$?

§3 Proof in Pred(V, $\mathscr{R}$)

To complete the construction of the logic called the First-Order Predicate Calculus on $(V, \mathscr{R})$, and henceforth denoted by Pred(V, $\mathscr{R}$), we have to define a proof in Pred(V, $\mathscr{R}$).

Definition 3.1. The *set of axioms* of Pred(V, $\mathscr{R}$) is the set $\mathscr{A} = \mathscr{A}_1 \cup \cdots \cup \mathscr{A}_5$, where

$$\mathscr{A}_1 = \{p \Rightarrow (q \Rightarrow p) | p, q \in P(V, \mathscr{R})\},$$

$$\mathscr{A}_2 = \{(p \Rightarrow (q \Rightarrow r)) \Rightarrow ((p \Rightarrow q) \Rightarrow (p \Rightarrow r)) | p, q, r \in P(V, \mathscr{R})\},$$

$$\mathscr{A}_3 = \{\sim \sim p \Rightarrow p | p \in P(V, \mathscr{R})\},$$

$\mathscr{A}_4 = \{(\forall x)(p \Rightarrow q) \Rightarrow (p \Rightarrow ((\forall x)q)) | p, q \in P(V, \mathscr{R}), x \notin \text{var}(p)\}$,
$\mathscr{A}_5 = \{(\forall x)p(x) \Rightarrow p(y) | p(x) \in P(V, \mathscr{R}), y \in V\}$.

We remind the reader that these axioms are stated in terms of elements of the reduced predicate algebra. In $\mathscr{A}_5$, for example, the substitution of y for x in $p(x)$ implies that we have chosen a representative of $[(\forall x)p(x)]$ in which $(\forall y)$ does not appear.

In addition to Modus Ponens, we shall use one further rule of inference, which will enable us to formalise the following commonly occurring argument: we have proved $p(x)$, but x was any element, and therefore $(\forall x)p(x)$. The rule of inference called Generalisation allows us to deduce $(\forall x)p(x)$ from $p(x)$ provided x is general. The restriction on the use of Generalisation needs to be stated carefully.

Definition 3.2 Let $A \subseteq P$, $p \in P$. A *proof of length* n of p from A is a sequence $p_1, \ldots, p_n$ of n elements of P such that $p_n = p$, the sequence $p_1, \ldots, p_{n-1}$ is a proof of length $n - 1$ of p_{n-1} from A, and

(a) $p_n \in \mathscr{A} \cup A$, or

(b) $p_i = p_j \Rightarrow p_n$ for some $i, j < n$, or

(c) $p_n = (\forall x)w(x)$ and some subsequence $p_{k_1}, \ldots, p_{k_r}$ of $p_1, \ldots p_{n-1}$ is a proof (of length $< n$) of $w(x)$ from a subset A_0 of A such that $x \notin \text{var}(A_0)$.

This is an inductive definition of a proof in Pred(V, $\mathscr{R}$). As for Prop(X), we require a proof to be a proof of finite length. The restriction $x \notin \text{var}(A_0)$ in (c) means that no special assumptions about x are used in proving $w(x)$, and is the formal analogue of the restriction on the use of Generalisation in our informal logic.

As before, we write $A \vdash p$ if there exists a proof of p from A. We denote by Ded(A) the set of all p such that $A \vdash p$. We write $\vdash p$ for $\varnothing \vdash p$, and any p for which $\vdash p$ is called a theorem of Pred(V, $\mathscr{R}$).

Example 3.3. We show $\{\sim(\exists x)(\sim p)\} \vdash (\forall x)p$ for any element $p \in P$. (Recall that $(\exists x)$ is an abbreviation for $\sim(\forall x)\sim$.) The following is a proof.

$p_1 = \sim\ \sim(\forall x)(\sim\ \sim p) \Rightarrow (\forall x)(\sim\ \sim p)$, ($\mathscr{A}_3$)

$p_2 = \sim\ \sim(\forall x)(\sim\ \sim p)$, (assumption)

$p_3 = (\forall x)(\sim\ \sim p)$, ($p_1 = p_2 \Rightarrow p_3$)

$p_4 = (\forall x)(\sim\ \sim p(x)) \Rightarrow\ \sim\ \sim p(y)$, ($\mathscr{A}_5$)

Note that by ($\mathscr{A}_5$), the y in p_4 may be chosen to be any variable. To permit a subsequent use of Generalisation, y must not be in var($\sim(\exists x)(\sim p(x))$). A possible choice for y is the variable x itself.

$p_5 = \sim\ \sim p(y)$, ($p_4 = p_3 \Rightarrow p_5$)

$p_6 = \sim\ \sim p(y) \Rightarrow p(y)$, ($\mathscr{A}_3$)

$p_7 = p(y)$, ($p_6 = p_5 \Rightarrow p_7$)

$p_8 = (\forall y)p(y)$. (Generalisation, $y \notin$ var($\sim(\exists x)(\sim p(x))$))

Exercises

3.4. Show that every axiom of $\text{Pred}(V, \mathscr{R})$ is valid.

3.5. Construct a proof in $\text{Pred}(V, \mathscr{R})$ of $(\forall x)(\forall y)p(x, y)$ from $\{(\forall y)(\forall x)p(x, y)\}$.

§4 Properties of Pred($V, \mathscr{R}$)

We have now constructed the logic $\text{Pred}(V, \mathscr{R})$. Its algebra of propositions is the reduced first order algebra $P(V, \mathscr{R})$, its valuations are the valuations associated with the interpretations of $P(V, \mathscr{R})$ defined in §2, and its proofs are as defined in §3.

We can immediately inquire if there is a substitution theorem for this logic, corresponding to Theorem 4.11 of the Propositional Calculus. There, substitution was defined in terms of a homomorphism $\varphi: P_1 \to P_2$ of one algebra of propositions into another. If P_1, P_2 are first order algebras, then as the concept of a homomorphism from P_1 to P_2 requires these algebras to have the same set of operations, it follows that they must have the same set of individual variables. Even in this case, a homomorphism would be too restrictive for our purposes, for we would naturally want to be able to interchange two variables x, y, so mapping elements $p(x)$ of the algebra to $\varphi(p(x)) = p(y)$, but unfortunately such a map is not a homomorphism. For if $p(x) \in P$ is such that $x \in \text{var}(p(x))$, $y \notin \text{var}(p(x))$, then

$$\varphi((\forall x)p(x)) = (\forall y)p(y) = (\forall x)p(x),$$
$$(\forall x)\varphi(p(x)) = (\forall x)p(y).$$

Since $y \in \text{var}((\forall x)p(y))$ but $y \notin \text{var}((\forall y)p(y))$, these elements are distinct and φ is not a homomorphism.

Definition 4.1. Let $P_1 = P(V_1, \mathscr{R}^{(1)})$ and $P_2 = P(V_2, \mathscr{R}^{(2)})$. A *semi-homomorphism* $(\alpha, \beta):(P_1, V_1) \to (P_2, V_2)$ is a pair of maps $\alpha: P_1 \to P_2$, $\beta: V_1 \to V_2$ such that

(a) $\beta(V_1)$ is infinite,

(b) α is an $\{F, \Rightarrow\}$-homomorphism, and

(c) $\alpha((\forall x)p) = (\forall x')\alpha(p)$, where $x' = \beta(x)$.

Lemma 4.2. *Let $(\alpha, \beta):(P_1, V_1) \to (P_2, V_2)$ be a semi-homomorphism. Let $p \in P_1$ and suppose $x \in V_1 - \text{var}(p)$. Then $\beta(x) \notin \text{var}(\alpha(p))$.*

Proof: We observe first that if $x \neq y$, then $(\forall x)p = (\forall y)p$ if and only if neither x nor y is in var(p).

Since $\beta(V_1)$ is infinite, there is an element $y' \in \beta(V_1)$ such that $y' \neq \beta(x)$ and $y' \notin \beta(\text{var}(p))$. Choosing $y \in V_1$ so that $\beta(y) = y'$, it follows that $(\forall x)p = (\forall y)p$. If $x' = \beta(x)$, then we have

$$(\forall x')\alpha(p) = \alpha((\forall x)p = \alpha((\forall y)p) = (\forall y')\alpha(p),$$

and it follows again that $x' \notin \text{var}(\alpha(p))$. □

Theorem 4.3. (The Substitution Theorem). *Let* $(\alpha, \beta):(P_1, V_1) \to (P_2, V_2)$ *be a semi-homomorphism. Let* $A \subseteq P$, $p \in P_1$.

(a) *If* $A \vdash p$, *then* $\alpha(A) \vdash \alpha(p)$.

(b) *If* $A \vDash p$, *then* $\alpha(A) \vDash \alpha(p)$.

Proof: (a) Let $p_1, \ldots, p_n$ be a proof of p from A. We use induction over n to show that $\alpha(p_1), \ldots, \alpha(p_n)$ is a proof of $\alpha(p)$ from $\alpha(A)$.

If $a = ((\forall x)(p \Rightarrow q)) \Rightarrow (p \Rightarrow (\forall x)q)$ is an axiom of type A_4, then by Lemma 4.2, the condition $x \notin \text{var}(p)$ is preserved by the semi-homomorphism (α, β), and so $\alpha(a)$ is again an axiom. In all other cases, it is clear that the image of an axiom is an axiom. Thus if $p \in \mathscr{A}^{(1)} \cup A$, then $\alpha(p) \in \mathscr{A}^{(2)} \cup \alpha(A)$, where $\mathscr{A}^{(i)}$ is the set of axioms of Pred$(V_i, \mathscr{R}^{(i)})$. Hence our desired result holds for $n = 1$.

For $n > 1$, we may suppose by induction that $\alpha(p_1), \ldots, \alpha(p_{n-1})$ is a proof of $\alpha(p_{n-1})$ from $\alpha(A)$. If $p_i = p_j \Rightarrow p_n$ for some $i, j < n$, then $\alpha(p_i) = \alpha(p_j) \Rightarrow \alpha(p_n)$, and the result holds. It remains only to consider the case that $p_n = (\forall x)q$, where some subsequence $q_1, \ldots, q_k$ of $p_1, \ldots, p_{n-1}$ is a proof of q from some subset $A_0 \subseteq A$ with $x \notin \text{var}(A_0)$. By induction, $\alpha(q_1), \ldots, \alpha(q_k)$ is a proof of $\alpha(q)$ from $\alpha(A_0)$. For each $w \in A_0$, $x \notin \text{var}(w)$, and by Lemma 4.2, $x' \notin \text{var}(\alpha(w))$, where $x' = \beta(x)$. Thus $x' \notin \text{var}(\alpha(A_0))$, and $\alpha(p_1), \ldots, \alpha(p_{n-1}), (\forall x')\alpha(q)$ is a proof. Since $(\forall x')\alpha(q) = \alpha((\forall x)q) = \alpha(p)$, the result is completely proved.

Part (b) is an easy consequence of (a) once we have proved the Adequacy Theorem, so we omit a proof. We leave as an exercise a direct proof of (b). □

Exercises

(The following exercises lead to a direct proof of part (b) of the Substitution Theorem. Throughout, $P_i = P(V_i, \mathscr{R}^{(i)})$ and $(\alpha, \beta):(P_1, V_1) \to (P_2, V_2)$ is a semi-homomorphism.)

4.4. Show that $(\forall x)p(x) = (\forall x)q(x)$ if and only if $p(x) = q(x)$.

4.5. We put $V_i^* = V_i \cup \{y\}$ and $P_i^* = P(V_i^*, \mathscr{R}^{(i)})$, where y is some new variable ($y \notin V_1 \cup V_2$). Show that for each $p(y) \in P_1^* - P_1$, there is a unique $q(y) \in P_2^*$ such that $\alpha((\forall x)p(x)) = (\forall x')q(x')$ for some $x \in V_1$, $x \notin \text{var}(p(y))$ and $x' = \beta(x)$. Hence show that there is a unique semi-homomorphism $(\alpha^*, \beta^*):(P_1^*, V_1^*) \to (P_2^*, V_2^*)$, extending (α, β), such that $\beta^*(y) = y$. Generalise to the addition of n new variables $y_1, \ldots, y_n$.

4.6. Let (U, φ, ψ, v) be an interpretation of P_2. For each $r \in \mathscr{R}_n^{(1)}$, we define an n-ary relation $\psi_1 r$ on U as follows. Take new variables $y_1, \ldots, y_n$, put $V_i^* = V_i \cup \{y_1, \ldots, y_n\}$, and construct the extension (α^*, β^*) of (α, β) as in 4.5. Given $(u_1, \ldots, u_n) \in U^n$, the mapping of y_i to u_i defines a unique extension of (U, φ, ψ, v) to P_2^*, and so assigns a value $v^*(q)$ to each $q \in P_2^*$. We define $(u_1, \ldots, u_n) \in \psi_1 r$ if and only if $v^*(\alpha^*(r(y_1, \ldots, y_n)) = 1$.

Show that $(U, \varphi\beta, \psi_1, v\alpha)$ is an interpretation of P_1. Hence prove part (b) of the Substitution Theorem.

Theorem 4.7. (The Soundness Theorem). *Let* $A \subseteq P(V, \mathscr{R})$, $p \in P(V, \mathscr{R})$. *If* $A \vdash p$, *then* $A \vDash p$.

Proof: Let $p_1, \ldots, p_n$ be a proof of p from A. Let (U, φ, ψ, v) be an interpretation of $P(V, \mathscr{R})$ such that $v(A) \subseteq \{1\}$. We have to show that $v(p) = 1$, and we shall use induction on n to prove it. If $n = 1$, $p \in \mathscr{A} \cup A$ and then $v(p) = 1$. Suppose by induction that $n > 1$ and the result holds for proofs of length less than n. If $p_i = p_j \Rightarrow p_n$ for some $i, j < n$, then $v(p_i) = v(p_j) = 1$, and it follows that $v(p) = 1$.

Suppose finally that $p_n = (\forall x)q(x)$ and that $q_1(x), \ldots, q_k(x)$ is a proof of $q(x)$ from the subset A_0 of A with $x \notin \text{var}(A_0)$. We must use condition (c_r) in the definition of interpretation, where r is the depth of p_n. Thus we take a new variable t, we put $V' = V \cup \{t\}$, and we consider extensions $\varphi': V' \to U$ of φ and maps $v'_{r-1}: P_{r-1}(V', \mathscr{R}) \to \mathbb{Z}_2$, as given in condition (c_r). We have to prove that in every case, $v'_{r-1}(q_k(t)) = 1$. But each v'_{r-1} extends uniquely to a valuation $v': P(V', \mathscr{R}) \to \mathbb{Z}_2$ such that (U, φ', ψ, v') is an interpretation of $P(V', \mathscr{R})$. By the Substitution Theorem (Theorem 4.3 (a)), $q_1(t), \ldots, q_k(t)$ is a proof of $q(t)$ from A_0, and so by induction (since $k < n$), $v'(q_k(t)) = 1$. Thus $v((\forall x)q(x)) = 1$ and the theorem is proved. □

Corollary 4.8. (The Consistency Theorem). *F is not a theorem of* Pred$(V, \mathscr{R})$.

Proof: Let U be any non-empty set, $\varphi: V \to U$ any function, and ψ any function on $\mathscr{R}$ such that if $r \in \mathscr{R}_n$, then $\psi(r)$ is an n-ary relation on U. Then there exists $v: P(V, \mathscr{R}) \to \mathbb{Z}_2$ such that (U, φ, ψ, v) is an interpretation. For every interpretation, and in particular for the one constructed above, $v(F) = 0$. The existence of one interpretation for which $v(F) = 0$ shows that F is not valid. The Soundness Theorem now shows that F is not a theorem. □

Theorem 4.9. (The Deduction Theorem). *Let $A \subseteq P = P(V, \mathscr{R})$ and let $p, q \in P$. Then $A \vdash p \Rightarrow q$ if and only if $A \cup \{p\} \vdash q$.*

Proof: If $A \vdash p \Rightarrow q$, then it follows, as in the case of the Propositional Calculus, that $A \cup \{p\} \vdash q$. Suppose $A \cup \{p\} \vdash q$. We shall again use induction over the length of the proof. The argument used for the case of the Propositional Calculus again applies except in the case where q is obtained by Generalisation. So we suppose $q = (\forall x)r(x)$ and $A_0 \vdash r(x)$, where $A_0 \subseteq A \cup \{p\}$ and $x \notin \text{var}(A_0)$.

(i) $p \notin A_0$. Then $A_0 \subseteq A$ and we have a proof of q from A_0. Follow this proof with the steps $q \Rightarrow (p \Rightarrow q)$, $p \Rightarrow q$ to obtain a proof of $p \Rightarrow q$ from A.

(ii) $p \in A_0$. We have a proof of $r(x)$ from A_0, and so by induction on the proof length, we have $A_1 \vdash p \Rightarrow r(x)$, where $A_1 = A_0 - \{p\}$. By Generalisation, a proof of $p \Rightarrow r(x)$ from A_1 may be followed with $(\forall x)(p \Rightarrow r(x))$. As $p \in A_0$ and $x \notin \text{var}(A_0)$, it follows that $x \notin \text{var}(p)$. We continue the proof with

$$(\forall x)(p \Rightarrow r(x)) \Rightarrow (p \Rightarrow (\forall x)r(x)) \qquad (\mathscr{A}_4)$$

and

$$p \Rightarrow (\forall x)r(x),$$

completing the proof and establishing the theorem. □

Example 4.10. As we did before, we use the techniques of the proof of the Deduction Theorem to convert the proof $\sim p$, $(\forall x)(\sim p)$, $\sim(\forall x)(\sim p)$, F of F from $\{(\exists x)p, \sim p\}$ ($x \notin \text{var}(p)$), into a proof of $\sim\ \sim p$ from $\{(\exists x)p\}$, so proving p from $\{(\exists x)p\}$.

Given proof	Comment	Corresponding Steps of Constructed Proof
$\sim p$	Assumption to be eliminated	$\sim p \Rightarrow ((\sim p \Rightarrow \sim p) \Rightarrow \sim p)$, $(\sim p \Rightarrow ((\sim p \Rightarrow \sim p) \Rightarrow \sim p)) \Rightarrow ((\sim p \Rightarrow (\sim p \Rightarrow \sim p)) \Rightarrow (\sim p \Rightarrow \sim p))$, $(\sim p \Rightarrow (\sim p \Rightarrow \sim p)) \Rightarrow (\sim p \Rightarrow \sim p)$, $\sim p \Rightarrow (\sim p \Rightarrow \sim p)$, $\sim p \Rightarrow \sim p$.
$(\forall x)(\sim p)$	Generalisation	$(\forall x)(\sim p \Rightarrow \sim p)$, $((\forall x)(\sim p \Rightarrow \sim p)) \Rightarrow (\sim p \Rightarrow (\forall x)(\sim p))$, $\sim p \Rightarrow (\forall x)(\sim p)$.
$\sim(\forall x)(\sim p)$	Retained assumption	$\sim(\forall x)(\sim p)$, $(\sim(\forall x)(\sim p)) \Rightarrow (\sim p \Rightarrow \sim(\forall x)(\sim p))$, $\sim p \Rightarrow (\sim(\forall x)(\sim p))$.
F	Modus ponens	$(\sim p \Rightarrow ((\forall x)(\sim p) \Rightarrow F)) \Rightarrow ((\sim p \Rightarrow (\forall x)(\sim p)) \Rightarrow (\sim p \Rightarrow F))$, $(\sim p \Rightarrow (\forall x)(\sim p)) \Rightarrow (\sim p \Rightarrow F)$, $\sim\ \sim p$.
	Extension to prove p	$\sim\ \sim p \Rightarrow p$, p.

Exercises

4.11. Convert the proof $(\forall x)p(x)$, $((\forall x)p(x)) \Rightarrow p(x)$, $p(x)$, $(\forall x)(p(x) \Rightarrow q(x))$, $(\forall x)(p(x) \Rightarrow q(x)) \Rightarrow (p(x) \Rightarrow q(x))$, $p(x) \Rightarrow q(x)$, $q(x)$, $(\forall x)q(x)$ of $(\forall x)q(x)$ from $\{(\forall x)(p(x) \Rightarrow q(x)), (\forall x)p(x)\}$ into a proof of $(\forall x)p(x) \Rightarrow (\forall x)q(x)$ from $\{(\forall x)(p(x) \Rightarrow q(x))\}$.

4.12. Prove $\{(\forall x)(p(x) \Rightarrow q(x)\} \vdash (\exists x)p(x) \Rightarrow (\exists x)q(x)$.

We now prove some lemmas which we shall need in establishing the Satisfiability Theorem. As for Prop(X), a subset A is consistent if $F \notin \text{Ded}(A)$.

Lemma 4.13. *Let A be a consistent subset of $P(V, \mathcal{R})$. Suppose $(\exists x)p(x) \in A$, and $t \notin \text{Var}(A)$. Then $F \notin \text{Ded}(A \cup \{p(t)\})$.*

Proof: Suppose $F \in \text{Ded}(A \cup \{p(t)\})$. Then by the Deduction Theorem, $\sim p(t) \in \text{Ded}(A)$. Since $t \notin \text{Var}(A)$, we may apply Generalisation and obtain $(\forall x)(\sim p(x)) \in \text{Ded}(A)$. But $(\exists x)p(x) = \sim(\forall x)(\sim p(x)) \in A$, and so $F \in \text{Ded}(A)$, contrary to assumption. □

Lemma 4.14. *Let A be a consistent subset of $P(V, \mathcal{R})$. Then there exist $V^* \supseteq V$ and $A^* \supseteq A$, where $A^* \subseteq P(V^*, \mathcal{R})$, such that*

(i) $F \notin \text{Ded}(A^*)$, *and*

(ii) *for all $p \in P(V^*, \mathcal{R})$, either $p \in A^*$ or $\sim p \in A^*$, and*

(iii) *if $(\exists x)p(x) \in A^*$, then for some $t \in V^*$, $p(t) \in A^*$.*

Proof: Put $V_0 = V$, $A_0 = A$, $P_0 = P(V, \mathscr{R})$. We construct inductively V_i, $P_i = P(V_i, \mathscr{R})$, A_i' and A_i for $i > 0$. Taking a new variable $t_p^{(i)}$ for each $p \in A_i$ of the form $p = (\exists x)q(x)$, we put

$$V_{i+1} = V_i \cup \{t_p^{(i)} | p \in A_i, p = (\exists x)q(x) \text{ for some } q(x)\},$$
$$A'_{i+1} = A_i \cup \{q(t_p^{(i)}) | p \in A_i, p = (\exists x)q(x), q(x) \in P_i\}.$$

Suppose that $F \notin \mathrm{Ded}(A_i)$. If $F \in \mathrm{Ded}(A'_{i+1})$, then $F \in \mathrm{Ded}(A_i \cup \{q_1(t_{p_1}^{(i)}), \cdots, q_r(t_{p_r}^{(i)})\})$ for some finite set $\{q_1(t_{p_1}^{(i)}), \ldots, q_r(t_{p_r}^{(i)})\}$, which is impossible by Lemma 4.13. Thus $F \notin \mathrm{Ded}(A'_{i+1})$, and by Lemma 2.11 of Chapter II, there exists $A_{i+1} \supseteq A'_{i+1}$ such that A_{i+1} satisfies (i) and (ii). For each $i > 0$, choose[2] such an A_i. Put $V^* = \bigcup_i V_i$, $A^* = \bigcup_i A_i$.

Since any finite subset of A^* is contained in some A_i, it follows that V^* and A^* satisfy (i), (ii) and (iii). □

Theorem 4.15. (The Satisfiability Theorem). *Let A be a consistent subset of $P(V, \mathscr{R})$. Then there exists an interpretation (U, φ, ψ, v) of $P(V, \mathscr{R})$ such that $v(A) \subseteq \{1\}$.*

Proof: If $V^* \supseteq V$ and $P(V^*, \mathscr{R}) \supseteq A^* \supseteq A$, then any interpretation of $P(V^*, \mathscr{R})$ for which $v(A^*) \subseteq \{1\}$ clearly restricts to an interpretation of $P(V, \mathscr{R})$ with $v(A) \subseteq \{1\}$. We may therefore suppose, without any loss of generality, that V, A satisfy the conditions (i), (ii) and (iii) of Lemma 4.14. To construct our interpretation, we take $U = V$, and $\varphi: V \to U$ the identity map. For each $r \in \mathscr{R}_n$, we put $\psi r = \{(x_1, \ldots, x_n) \in V^n | r(x_1, \ldots, x_n) \in A\}$. For each $p \in P(V, \mathscr{R})$, we put $v(p) = 1$ if $p \in A$ and $v(p) = 0$ otherwise. It is easily checked that (U, φ, ψ, v) satisfies the conditions (a), (b) of the definition of an interpretation, and we are left with showing that the condition (c_k) is satisfied for all k.

Let t be some new variable, and let $p = (\forall x)q(x)$ have depth $k + 1$. Suppose first that $p \in A$. Let φ' be any extension of φ to $V' = V \cup \{t\}$, and let $v_k': P_k(V', \mathscr{R}) \to Z_2$ be as required for condition (c_{k+1}). Put $y = \varphi'(t)$. Since, by induction, v satisfies (c_i) for $i \leqslant k$, it follows that for all $w(x) \in P_k$, $v'(w(t)) = v(w(y))$. Now $(\forall x)q(x) \in A$, therefore $q(y) \in \mathrm{Ded}(A) = A$, since A is a maximal consistent subset, and this holds for all $y \in V$. Thus $v'(q(t)) = v(q(y)) = 1$ and condition (c_{k+1}) is satisfied in this case.

Suppose that $p = (\forall x)q(x) \notin A$. As $\{\sim(\exists x)(\sim q(x))\} \vdash (\forall x)q(x)$, it follows that $\sim(\exists x)(\sim q(x)) \notin A$. Hence $(\exists x)(\sim q(x)) \in A$, and so for some $y \in V$, $\sim q(y) \in A$. Consider the extension φ' of φ to V' with $\varphi'(t) = y$, and the corresponding $v_k': P(V', \mathscr{R}) \to Z_2$. Then $v'(q(t)) = v(q(y)) = 0$. As $v(p) = 0$, we see again that condition (c_{k+1}) is satisfied. □

Theorem 4.16. (The Adequacy Theorem). *Let $A \subseteq P(V, \mathscr{R})$, $p \in P(V, \mathscr{R})$. If $A \vDash p$, then $A \vdash p$.*

[2] The proof of Lemma 2.11 involved an application of Zorn's Lemma. We also use the (countable) axiom of choice here to select the A_i.

Proof: If $F \notin \operatorname{Ded}(A \cup \{\sim p\})$, then by the Satisfiability Theorem, there exists an interpretation (U, φ, ψ, v) of $P(V, \mathscr{R})$ such that $v(A \cup \{\sim p\}) \subseteq \{1\}$, which contradicts the hypothesis $A \models p$. Therefore $A \cup \{\sim p\} \vdash F$. Hence, by the Deduction Theorem, $A \vdash \sim \sim p$, and the result follows. □

Corollary 4.17. (The Compactness Theorem). *If $A \models p$, then $A_0 \models p$ for some finite subset A_0 of A.*

The Soundness Theorem and the Adequacy Theorem together show that if $A \subseteq P(V, \mathscr{R})$ and $p \in P(V, \mathscr{R})$, then $A \models p$ if and only if $A \vdash p$. This result is usually called Gödel's (or the Gödel-Henkin) Completeness Theorem. It was first proved by Gödel in 1930. The method of proof we have used, depending on the Satisfiability Theorem, is due to Henkin.

We have now established for Pred($V, \mathscr{R}$) all the properties previously established for Prop(X), with the exception of decidability. We have good reason for not attempting to prove Pred($V, \mathscr{R}$) is decidable. If $\mathscr{R}$ contains at least one relation symbol of arity greater than 1, then Pred($V, \mathscr{R}$) is undecidable. The precise meaning of this statement, and its proof (which is due to Church and Kalmar), are given in Chapter IX.

Exercise 4.18. An element $p \in P(V, \mathscr{R})$ is said to be expressed in prenex normal form when it is expressed in the form $p = Q_1 Q_2 \cdots Q_k q$, where Q_i is either $(\forall x_i)$ or $(\exists x_i)$, $x_1, \ldots, x_k$ are distinct, and q is a quantifier-free element of $P(V, \mathscr{R})$. Give an algorithm which constructs from any $p \in P(V, \mathscr{R})$, an element p' in prenex normal form such that $\vdash (p \Rightarrow p') \wedge (p' \Rightarrow p)$.

Chapter V

First-Order Mathematics

§1 Predicate Calculus with Identity

In this chapter, we shall reconstruct some parts of ordinary mathematics within the logical system constructed in Chapter IV. A piece of mathematics constructed within the first-order predicate calculus will be called a first-order theory. By comparing a first-order theory with the informal theory on which it is modelled, we may gain insight into the influence of our logical system on our mathematics.

One feature common to all mathematical theories is the concept of equality or identity. A statement of the form $a = b$ always means that a and b denote the same mathematical object. A consequence of $a = b$ is that, in any statement involving a, we may replace any of the occurrences of a by b without altering the truth or falsity of the statement. We therefore begin by investigating how to formalise in $\mathrm{Pred}(V, \mathscr{R})$ the concept of identity. We clearly require a binary relation symbol $\mathscr{I} \in \mathscr{R}_2$. As the axioms of identity, we take the set $I \subseteq P(V, \mathscr{R})$ consisting of $(\forall x)\mathscr{I}(x, x)$ and the elements $(\forall x_1)\cdots(\forall x_n)(\forall y)(\mathscr{I}(x_j, y) \Rightarrow (r(x_1, \ldots, x_n) \Rightarrow r(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)))$, for all $r \in \mathscr{R}_n$, all n, and all $j \leqslant n$.

Exercises

1.1. Prove $I \vdash \mathscr{I}(x, y) \Rightarrow \mathscr{I}(y, x)$.

1.2. Prove $I \vdash \mathscr{I}(x, y) \Rightarrow (\mathscr{I}(y, z) \Rightarrow \mathscr{I}(x, z))$.

1.3. Let $w(x, z)$ be any element of P, possibly involving other variables besides x, z. Show that $I \vdash \mathscr{I}(x, y) \Rightarrow (w(x, x) \Rightarrow w(y, x))$. (Hint: use induction over the number of steps in the construction of $w(x, y)$ from V and $\mathscr{R}$.)

1.4. Let (U, φ, ψ, v) be an interpretation of $P(V, \mathscr{R})$ such that $\psi\mathscr{I}$ is the identity relation on U. Let U' be any set containing U, and let $\pi: U' \to U$ be any function such that $\pi(u) = u$ for all $u \in U$. Let $\varphi': V \to U'$ be the composition of φ with the inclusion map $U \to U'$. For each $r \in \mathscr{R}_n$, define the n-ary relation $\psi'r$ on U' by $(u'_1, \ldots, u'_n) \in \psi'r$ if and only if $(\pi(u'_1), \ldots, \pi(u'_n)) \in \psi r$. Show that this defines an interpretation $(U', \varphi', \psi', v')$ of $P(V, \mathscr{R})$, and that for $p \in P(V, \mathscr{R})$, we have $v'(p) = v(p)$. Show that $\psi'\mathscr{I}$ is an equivalence relation on U', but that, no matter what the interpretation (U, φ, ψ, v), U' and π can be constructed such that $\psi'\mathscr{I}$ is not the relation of identity in U'.

According to Exercise 1.4, no matter what subset $I' \supseteq I$ of $P(V, \mathscr{R})$ we choose as our axioms of identity, we cannot thereby force $\psi\mathscr{I}$ to be the relation of identity in every interpretation of $P(V, \mathscr{R})$ such that $v(I') = \{1\}$, unless of course we have $F \in \mathrm{Ded}(I')$ and so exclude the existence of such interpretations. We overcome this by constructing a modified form of the first-order predicate calculus, in which the only interpretations allowed will be those for which $\psi\mathscr{I}$ is the identity relation.

Definition 1.5. Suppose $\mathscr{I} \in \mathscr{R}_2$. A *proper interpretation* of $P(V, \mathscr{R})$ is an interpretation (U, φ, ψ, v) such that $\psi\mathscr{I}$ is the relation of identity on U.

Definition 1.6. $\mathrm{Pred}_{\mathscr{I}}(V, \mathscr{R})$ is the logic with algebra of propositions $P(V, \mathscr{R} \cup \{\mathscr{I}\})$, valuations those arising from proper interpretations, and with proof of p from A in $\mathrm{Pred}_{\mathscr{I}}(V, \mathscr{R})$ defined as a proof of p from $I \cup A$ in $\mathrm{Pred}(V, \mathscr{R} \cup \{\mathscr{I}\})$.

We shall always assume $\mathscr{I} \in \mathscr{R}$, and so have $P(V, \mathscr{R})$ as the algebra of propositions. We write $A \vdash_{\mathscr{I}} p$ and $p \in \mathrm{Ded}_{\mathscr{I}}(A)$ to indicate that p is provable from A in $\mathrm{Pred}_{\mathscr{I}}(V, \mathscr{R})$, i.e., that $A \cup I \vdash p$ or equivalently $p \in \mathrm{Ded}(A \cup I)$. We say that p is a *proper consequence* of A, written $A \models_{\mathscr{I}} p$ or $p \in \mathrm{Con}_{\mathscr{I}}(A)$, if $v(p) = 1$ for every proper interpretation of $P(V, \mathscr{R})$ with $v(A) \subseteq \{1\}$. Because of the restriction on the interpretations considered, $A \models_{\mathscr{I}} p$ would appear to be weaker than $A \cup I \models p$. We shall see shortly that they are in fact equivalent.

Theorem 1.7. (The Satisfiability Theorem) *Suppose* $F \notin \mathrm{Ded}_{\mathscr{I}}(A)$. *Then there exists a proper interpretation of* $P(V, \mathscr{R})$ *with* $v(A) \subseteq \{1\}$.

Proof: Since $F \notin \mathrm{Ded}(A \cup I)$, there exists an interpretation (U, φ, ψ, v) of $P = P(V, \mathscr{R})$ such that $v(A \cup I) = \{1\}$. The relation $\psi\mathscr{I}$ is an equivalence relation on U. For $u \in U$, denote by $\bar{u}$ the equivalence class $\{u' \in U \mid (u, u') \in \psi\mathscr{I}\}$, and let $\bar{U}$ be the set of all these equivalence classes. Define $\bar{\varphi}: V \to \bar{U}$ by $\bar{\varphi}(x) = \overline{\varphi(x)}$ for all $x \in V$. For each $r \in \mathscr{R}_n$, ψr has the property that if $(u_i, u_i') \in \psi\mathscr{I}$, then $(u_1, \ldots, u_n) \in \psi r$ if and only if $(u_1', \ldots, u_n') \in \psi r$. Hence we can define a relation $\bar{\psi} r$ on $\bar{U}$ by putting $(\bar{u}_1, \ldots, \bar{u}_n) \in \bar{\psi} r$ if and only if $(u_1, \ldots, u_n) \in \psi r$. This defines a function $\bar{\psi}$ from $\mathscr{R}$ into the relations on $\bar{U}$, and it is easily checked that $(\bar{U}, \bar{\varphi}, \bar{\psi}, v)$ is a proper interpretation of $P(V, \mathscr{R})$. The valuation v is unchanged, consequently we have a proper interpretation with $v(A) \subseteq \{1\}$. □

Corollary 1.8.

(i) $\mathrm{Con}_{\mathscr{I}}(A) = \mathrm{Con}(A \cup I)$

(ii) *If* $A \models_{\mathscr{I}} p$, *then* $A \vdash_{\mathscr{I}} p$.

The soundness and consistency of $\mathrm{Pred}_{\mathscr{I}}(V, \mathscr{R})$ both follow immediately from the corresponding properties of $\mathrm{Pred}(V, \mathscr{R})$.

§2 First-Order Mathematical Theories

A branch of mathematics is defined by listing the properties and relationships to be studied and by listing the assumptions (usually known as the axioms of the branch of mathematics) made about them. For example, in plane projective geometry, the only properties considered are those of being called a point or line (the actual nature of the objects is irrelevant, only the way they are divided into the two classes matters) and we are concerned with the one relationship of a point lying on a line. (It is taken for granted that we also use the relationship of identity.) The axioms of plane projective geometry are that through any two distinct points there is one and only one line, that any two distinct lines have one and only one common point, and the non-triviality axiom that there exist four points such that no three of them are collinear.

We shall define a mathematical theory in terms of lists of relations and axioms. It is convenient also to include a list of any special objects named in the axioms.

Definition 2.1. A *first-order mathematical theory* is a triple $\mathscr{T} = (\mathscr{R}, A, C)$ where $\mathscr{I} \in \mathscr{R}$, $A \subseteq P(V, \mathscr{R})$ for some $V \supset C$ such that $V - C$ is infinite, and $\text{var}(A) = C$. The set A is called the set of (*mathematical*) *axioms* of $\mathscr{T}$, the set C is called the set of (*individual*) *constants* of $\mathscr{T}$, while the *language*[1] of $\mathscr{T}$ is the subset $\mathscr{L}(\mathscr{T}) = \{p \in P(V, \mathscr{R}) | \text{var}(p) \subseteq C\}$ of $P(V, \mathscr{R})$. A *theorem* of $\mathscr{T}$ is an element $p \in \mathscr{L}(\mathscr{T})$ such that $A \vdash_{\mathscr{I}} p$.

We point out that the set V is not specified in $\mathscr{T}$, and that any suitable set V may be taken. The set $\mathscr{L}(\mathscr{T})$ is independent of the choice of V. Later, we shall occasionally need a standardised set V of variables, such that $V - C$ is countably infinite. We select as standard variable set the set $V_0 = C \cup \{x_i | i \in \mathbf{N}\}$, where the x_i are disjoint from C.

Definition 2.2. The *algebra* of $\mathscr{T}$ is the set $P(\mathscr{T}) = P(V_0, \mathscr{R})$, where V_0 is the standard variable set. An element $p \in P(\mathscr{T})$, such that $\text{var}(p) \subseteq \{x_1, \ldots, x_n\} \cup C$, is called an *n-variable formula* of $\mathscr{T}$.

The following notations will be used in discussing first-order theories $\mathscr{T}$. If $U \subseteq P(V, \mathscr{R})$ and $p \in P(V, \mathscr{R})$, then we write $U \vdash_{\mathscr{T}} p$ for $A \cup U \vdash_{\mathscr{I}} p$, $\mathscr{T} \vdash p$(or $\vdash_{\mathscr{T}} p$) for $A \vdash_{\mathscr{I}} p$, and $U \vDash_{\mathscr{T}} p$, $\mathscr{T} \vDash_{\mathscr{I}} p$(or $\vDash_{\mathscr{T}} p$), for $A \cup U \vdash_{\mathscr{I}} p$ and $A \vDash_{\mathscr{I}} p$ respectively.

Examples

2.3. (First-order plane projective geometry) We take two unary predicate symbols p, ℓ, interpreting $p(x)$ as "x is a point", and $\ell(x)$ as "x is a line". We take a binary predicate symbol $\in$, and interpret $\in(x, y)$ as "x lies on y". These express the basic concepts of plane projective geometry, so we take $\mathscr{R} = \{p, \ell, \in, \mathscr{I}\}$. Our axiom set is the set $\mathscr{A} = \{a_1, \ldots, a_6\}$, where

$$a_1 = (\forall x)((p(x) \vee \ell(x)) \wedge \sim (p(x) \wedge \ell(x))),$$

[1] The reader is warned that most authors use this term for $P(V, \mathscr{R})$.

$$\begin{aligned}
a_2 &= (\forall x)((\exists y) \in(x, y) \Rightarrow p(x)),\\
a_3 &= (\forall x)((\exists y) \in(y, x) \Rightarrow \ell(x)),\\
a_4 &= (\forall x)(\forall y)(p(x) \wedge p(y) \wedge \sim\mathscr{I}(x, y) \Rightarrow (\exists z)(\in(x, z) \wedge \in(y, z)\\
&\qquad \wedge (\forall t)(\in(x, t) \wedge \in(y, t) \Rightarrow \mathscr{I}(z, t)))),\\
a_5 &= (\forall x)(\forall y)(\ell(x) \wedge \ell(y) \wedge \sim\mathscr{I}(x, y) \Rightarrow (\exists z)(\in(z, x) \wedge \in(z, y)\\
&\qquad \wedge (\forall t)(\in(t, x) \wedge \in(t, y) \Rightarrow \mathscr{I}(z, t)))),\\
a_6 &= (\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(p(x_1) \wedge p(x_2) \wedge p(x_3) \wedge p(x_4) \wedge \sim\mathscr{I}(x_1, x_2) \wedge\\
&\qquad \sim\mathscr{I}(x_1, x_3) \wedge \sim\mathscr{I}(x_1, x_4) \wedge \sim\mathscr{I}(x_2, x_3) \wedge \sim\mathscr{I}(x_2, x_4) \wedge \sim\mathscr{I}(x_3, x_4)\\
&\qquad \wedge \sim c(x_1, x_2, x_3) \wedge \sim c(x_1, x_2, x_4) \wedge \sim c(x_1, x_3, x_4) \wedge \sim c(x_2, x_3, x_4))
\end{aligned}$$

where in the non-triviality axiom a_6, $c(x_1, x_2, x_3)$ denotes $(\exists z)(\in(x_1, z) \wedge \in(x_2, z) \wedge \in(x_3, z))$. The axiom a_1 says that each object is either a point or a line, but not both. Axioms a_2 and a_3 say that $\in$ is a relation between a point and a line, while axioms a_4 and a_5 are the usual incidence axioms. For this theory, the set $C = \varnothing$.

There is a very useful notation which abbreviates axioms such as a_4 and a_5. We write $(\exists! x)w(x)$ for $(\exists x)(w(x) \wedge (\forall y)(w(y) \Rightarrow \mathscr{I}(x, y)))$, where $w(x)$ is any element of $P(V, \mathscr{R})$. $(\exists! x)w(x)$ may be read "There exists a unique x such that $w(x)$". In this notation, we have

$$a_4 = (\forall x)(\forall y)(p(x) \wedge p(y) \wedge \sim\mathscr{I}(x, y) \Rightarrow (\exists! z)(\in(x, z) \wedge \in(y, z))).$$

2.4. (Elementary group theory) We take $\mathscr{R} = \{\mathscr{I}, m\}$, where m is a ternary relation symbol. We interpret $m(x, y, z)$ as "$xy = z$". For axioms, we take $A = \{a_1, \ldots, a_4\}$, where

$$\begin{aligned}
a_1 &= (\forall x)(\forall y)(\exists! z)m(x, y, z),\\
a_2 &= (\forall x)(\forall y)(\forall z)(\forall a)(\forall b)(\forall c)(\forall d)(m(x, y, a) \wedge m(a, z, b)\\
&\qquad \wedge m(y, z, c) \wedge m(x, c, d) \Rightarrow \mathscr{I}(b, d)),\\
a_3 &= (\forall x)m(e, x, x),\\
a_4 &= (\forall x)(\exists y)m(y, x, e).
\end{aligned}$$

Axiom a_1 asserts that m defines a function, a_2 is the associative law, a_3 asserts that e is a left identity and a_4 asserts the existence of left inverses. We have $C = \{e\}$. We could reformulate the theory without individual constants by replacing a_3 and a_4 by $(\exists e)(a_3 \wedge a_4)$.

This theory is too restrictive for the study of group theory. Within it, we can prove results such as that e is a right identity or that the identity is unique. But we have no way of expressing properties of subsets, so we cannot discuss subgroups. Nor can we discuss relationships between groups. We called this theory *elementary* group theory because it is restricted to the relationships between elements of a group (as distinct from the relationships between subsets of a group). We shall use the word "elementary" with this meaning in relation to other theories.

Exercises

2.5. Show that $(\forall x)m(x, e, x)$, $(\forall e')((\forall x)m(e', x, x) \Rightarrow \mathscr{I}(e, e'))$, $(\forall x)(\forall y)(m(y, x, e) \Rightarrow m(x, y, e))$ are theorems of the elementary group theory of Example 2.4.

2.6. Show that the formal analogue of the statement "There exist four distinct lines, no three of which are concurrent" is a theorem of the first-order plane projective geometry of Example 2.3.

2.7. $\mathscr{T} = (\mathscr{R}, A, C)$ is a first-order theory and $(\alpha, \beta):(P(V, \mathscr{R}), V) \to (P(V, \mathscr{R}), V)$ is a semi-homomorphism such that $\alpha(A) \subseteq \mathrm{Ded}_f(A)$. If $\mathscr{T} \vdash p$, prove that $\mathscr{T} \vdash \alpha(p)$.

2.8. $\mathscr{T}$ is the first-order plane projective geometry of Example 2.3. The dual $\bar{w}$ of an element $w \in P(V, \mathscr{R})$ is the element obtained from w by interchanging p and ℓ and replacing $\in(x, y)$ by $\in(y, x)$ (all $x, y \in V$) throughout Show that if α is the map $\alpha(w) = \bar{w}$, and if β is the identity map, then (α, β) is a semi-homomorphism (in fact an automorphism) of $P(V, \mathscr{R})$, satisfying the condition of Exercise 2.7. Hence prove that the dual of a theorem of $\mathscr{T}$ is a theorem of $\mathscr{T}$.

The examples given above show how particular mathematical systems may be used to construct first-order theories. We regard the concept of a first-order theory as fundamental to our study of the relationship between reasoning and mathematics, and our direction is set firmly by the next definition. We denote by $\mathrm{rel}(M)$ the set of all relations on a set M.

Definition 2.9. A *model* of the first-order theory $\mathscr{T} = (\mathscr{R}, A, C)$ is a set M together with functions $v:C \to M$, $\psi:\mathscr{R} \to \mathrm{rel}(M)$, such that for some set V of variables ($V \supset C$, $V - C$ infinite), there exists a proper interpretation (M, φ, ψ, v) of $P(V, \mathscr{R})$ for which $\varphi|_C = v$ and $v(A) \subseteq \{1\}$.

We think of a model (M, v, ψ) of the theory $\mathscr{T}$ as the essential part of a proper interpretation of $\mathscr{T}$ for which the axioms of $\mathscr{T}$ are true (i.e., for which $v(A) \subseteq \{1\}$). Although the valuation v of $P(V, \mathscr{R})$ is determined by φ and ψ, the restriction $v|_{\mathscr{L}(\mathscr{T})}$ is completely determined by v and ψ, and is independent of the choice of V and the interpretation. Hence there is a well-determined valuation v of $\mathscr{L}(\mathscr{T})$ corresponding to each model (M, v, ψ) of $\mathscr{T}$, and we say that $p \in \mathscr{L}(\mathscr{T})$ is *true for the model* (M, v, ψ) of $\mathscr{T}$ if $v(p) = 1$. We shall refer to the model (M, v, ψ) of $\mathscr{T}$ as the model M of $\mathscr{T}$, whenever this abuse of notation does not lead to confusion.

Example 2.10. Let G be a group with multiplication written as juxtaposition and with identity element 1. We put $v(e) = 1$, and we put $\psi m = \{(x, y, z) \in G^3 | xy = z\}$. $\psi\mathscr{I}$ will of course be the identity relation. Then $G = (G, v, \psi)$ is a model of the elementary group theory of Example 2.4. A model of the elementary group theory is essentially a group.

Given a model (M, v, ψ) of the theory $\mathscr{T}$, some relations on the set M are derived naturally from $\mathscr{T}$, in the following manner. Let $p(x_1, \ldots, x_n)$ be an

n-variable formula of $\mathcal{T}$. For any $m_1, \ldots, m_n \in M$, there is an interpretation (M, φ, ψ, v) of $P(\mathcal{T})$ such that $\varphi|_C = v$ and $\varphi(x_i) = m_i$ for $i = 1, \ldots, n$. These conditions on φ determine $v(p)$, and if $v(p) = 1$, we say that $(m_1, \ldots, m_n)$ satisfies $p(x_1, \ldots, x_n)$, or (by abuse of language) that $p(m_1, \ldots, m_n)$ is true in M. Hence $p(x_1, \ldots, x_n)$ defines an n-ary relation on M, which (by abuse of notation) we denote by $\psi(p)$:

$$\psi(p) = \{(m_1, \ldots, m_n) \in M^n | p(m_1, \ldots, m_n) \text{ is true in } M\}.$$

This leads to the following definition.

Definition 2.11. The n-ary relation ρ on the model M of $\mathcal{T}$ is said to be *definable in* $\mathcal{T}$ if $\rho = \psi(p)$ for some n-variable formula p of $\mathcal{T}$. The function $f: M^n \to M$ is called a *definable function* if there is an $(n + 1)$-variable formula p of $\mathcal{T}$ such that

(i) for all $a_1, \ldots, a_n, b \in M, f(a_1, \ldots, a_n) = b$ if and only if $p(a_1, \ldots, a_n, b)$ is true, and

(ii) $\mathcal{T} \vdash (\forall x_1) \cdots (\forall x_n)(\exists! y)p(x_1, \ldots, x_n, y)$.

Example 2.12. Conjugacy is a definable relation in elementary group theory. It is defined by the formula

$$p(x_1, x_2) = (\exists x_3)(\exists x_4)(\exists x_5)(m(x_3, x_4, e) \wedge m(x_3, x_1, x_5) \wedge m(x_5, x_4, x_2)).$$

Inverse is a definable function, defined by the formula

$$q(x_1, x_2) = m(x_2, x_1, e).$$

§3 Properties of First-Order Theories

Definition 3.1. The first-order theory $\mathcal{T}$ is called *consistent* if F is not a theorem of $\mathcal{T}$.

The Soundness and Satisfiability Theorems for $\text{Pred}_{\mathcal{I}}(V, \mathcal{R})$ immediately give the following result.

Theorem 3.2. *The theory $\mathcal{T}$ is consistent if and only if there exists a model of $\mathcal{T}$.*

Definition 3.3. The theory $\mathcal{T}$ is called *complete* if, for every $p \in \mathcal{L}(\mathcal{T})$, either $\mathcal{T} \vdash p$ or $\mathcal{T} \vdash {\sim}p$.

Elementary group theory is not complete, for consider $p = (\forall x)\mathcal{I}(x, e) \in \mathcal{L}(\mathcal{T})$. If p were a theorem of the theory, it would be true for every model. But p is true for the group G if and only if the order of G is 1. As there are groups of order greater than 1, p cannot be a theorem. As there are groups of order 1, ${\sim}p$ cannot be a theorem.

Theorem 3.4. *The first-order theory $\mathcal{T}$ is complete if and only if every $p \in \mathcal{L}(\mathcal{T})$ which is true in one model of $\mathcal{T}$ is true in every model of $\mathcal{T}$.*

Proof: The result is trivial if $\mathscr{T}$ is inconsistent, so we suppose $\mathscr{T}$ consistent. Suppose that $\mathscr{T}$ is complete, and that $p \in \mathscr{L}(\mathscr{T})$ is true in the model M. Since $\sim p$ is false for M, it is not a theorem of $\mathscr{T}$, and, since $\mathscr{T}$ is complete, it follows that $\mathscr{T} \vdash p$. Therefore p is true in every model of $\mathscr{T}$.

Suppose conversely that for all $p \in \mathscr{L}(\mathscr{T})$, p true in one model imples p true in every model. Take some model M of $\mathscr{T}$, and let $p \in \mathscr{L}(\mathscr{T})$. If p is true in M, then p is true in every model, i.e., $\mathscr{T} \models p$, and hence by the Adequacy Theorem $\mathscr{T} \vdash p$. If p is false in M, then $\sim p$ is true in M and so $\mathscr{T} \vdash \sim p$. Thus $\mathscr{T}$ is complete. □

Examples of complete theories are easily produced, as we now show.

Theorem 3.5. *Let $\mathscr{T} = (\mathscr{R}, A, C)$ be a consistent theory. Then there exists $A' \subseteq \mathscr{L}(\mathscr{T})$, with $A' \supseteq A$ and such that $(\mathscr{R}, A', C)$ is consistent and complete.*

Proof: Since $\mathscr{T}$ is consistent, it has a model M say. Put $A' = \{p \in \mathscr{L}(\mathscr{T}) | p$ true in $M\}$. Then A' has the required properties. □

Definition 3.6. Let (M_1, v_1, ψ_1) and (M_2, v_2, ψ_2) be models of the theory $\mathscr{T}$. We say that M_1 is *isomorphic* to M_2 if there exists a bijective map $\alpha: M_1 \to M_2$ such that $\alpha v_1 = v_2$ and $(m_1, \ldots, m_n) \in \psi_1 r$ if and only if $(\alpha(m_1), \ldots, \alpha(m_n)) \in \psi_2 r$ for all $r \in \mathscr{R}_n$, all $m_1, \ldots, m_n \in M_1$, and all $n \in \mathbf{N}$.

Definition 3.7. The theory $\mathscr{T}$ is called *categorical* if all models of $\mathscr{T}$ are isomorphic.

Examples

3.8. Two models G_1, G_2 of elementary group theory are isomorphic as models if and only if they are isomorphic in the group theoretic sense. Since there exist groups G_1, G_2 which are not isomorphic, elementary group theory is not categorical.

3.9. We form trivial group theory by adding the further axiom $(\forall x)\mathscr{I}(x, e)$ to elementary group theory. A model of trivial group theory is a group of order 1. Any two such groups are isomorphic, thus trivial group theory is categorical.

Observe that if M_1, M_2 are isomorphic models of the theory $\mathscr{T}$ and if $p \in \mathscr{L}(\mathscr{T})$ is true for M_1, then it is true for M_2. The definition of categoricity, together with Theorem 3.4, immediately yields the next theorem.

Theorem 3.10. *If the theory $\mathscr{T}$ is categorical, then $\mathscr{T}$ is complete.*

We now generalise the concept of categoricity. We shall denote the cardinal of any set X by $|X|$.

Definition 3.11. The *cardinal* of a model $M = (M, v, \psi)$ is the cardinal $|M|$ of the set M, and will be denoted by $|M|$.

Note that isomorphic models have the same cardinal.

Definition 3.12. Let χ be a cardinal number. The theory $\mathscr{T}$ is called *χ-categorical* or *categorical in cardinal* χ if all models of $\mathscr{T}$ which have cardinal χ are isomorphic.

Example 3.13. Elementary group theory is categorical in cardinal 1. It is not categorical in cardinal 4, because there are two distinct isomorphism classes of groups of order 4.

Provided that χ is a finite cardinal, there is in the language $\mathscr{L}(\mathscr{T})$ of any theory $\mathscr{T}$ an element which specifies χ as the cardinal of a model of $\mathscr{T}$. For if we denote by al(n) the proposition

$$(\exists a_1)\cdots(\exists a_n)(\sim\mathscr{I}(a_1, a_2) \wedge \sim\mathscr{I}(a_1, a_3) \wedge \cdots \wedge \sim\mathscr{I}(a_1, a_n)$$
$$\wedge \sim\mathscr{I}(a_2, a_3) \wedge \cdots \wedge \sim\mathscr{I}(a_{n-1}, a_n)),$$

then any model of $\mathscr{T}$ in which al(n) is true has at least n elements. Any model in which al(n) $\wedge \sim$al($n + 1$) is true has exactly n elements.

Theorem 3.14. *Suppose the theory $\mathscr{T}$ has models of arbitrarily large finite cardinal. Then $\mathscr{T}$ has an infinite model.*

Proof: Let $\mathscr{T} = (\mathscr{R}, A, C)$, and put $\mathscr{T}' = (\mathscr{R}, A', C)$ where $A' = A \cup \{\text{al}(n) | n \in \mathrm{N}^+\}$. We show that $\mathscr{T}'$ is consistent. If $A' \vdash_{\mathscr{I}} F$, then $A \cup N \vdash_{\mathscr{I}} F$ for some finite subset N of $\{\text{al}(n) | n \in \mathrm{N}^+\}$. Let $n_0 = \max\{n | \text{al}(n) \in N\}$. By hypothesis, there exists a model M of $\mathscr{T}$ with $|M| \geqslant n_0$. This M is a model of $(\mathscr{R}, A \cup N, C)$, which contradicts the hypothesis $A \cup N \vdash_{\mathscr{I}} F$. Hence $\mathscr{T}'$ is consistent, and so it has a model. Any model M of $\mathscr{T}'$ must satisfy $|M| \geqslant n$ for all $n \in \mathrm{N}^+$, hence $|M|$ is infinite. □

Exercises

3.15. R is a ring with 1. Construct an elementary theory (i.e., one concerned with elements and not with subsets or maps) Mod_R, of unital (left) R-modules, such that the models of the theory are precisely all unital R-modules. (Hint: take each $r \in R$ as a binary relation symbol, interpreting $r(m_1, m_2)$ as $rm_1 = m_2$.)

3.16. Construct an elementary theory of fields with constants 0, 1. In the language $\mathscr{L}(\mathscr{F})$ of this theory $\mathscr{F}$, construct a proposition char(n), which asserts that the characteristic divides n ($n \in \mathrm{N}^+$). Hence construct a theory $\mathscr{F}_0$ of fields of characteristic 0 such that $\mathscr{L}(\mathscr{F}_0) = \mathscr{L}(\mathscr{F})$ and the set A_0 of axioms of $\mathscr{F}_0$ includes the set A of axioms of $\mathscr{F}$. Show that for each theorem p of $\mathscr{F}_0$, there is a number $n_p \in \mathrm{N}$ such that p is true for all fields of characteristic greater than n_p. Show that no set A_1 of axioms, such that $\mathscr{L}(\mathscr{F}) \supseteq A_1 \supseteq A$ and A_1 contains only finitely many elements of $\mathscr{L}(\mathscr{F}) - A$, can axiomatise fields of characteristic 0.

Definition 3.17. The *cardinal* $|\mathscr{T}|$ of the theory $\mathscr{T} = (\mathscr{R}, A, C)$ is $|\mathscr{R} \cup A|$. $\mathscr{T}$ is called *finite* if $\mathscr{R} \cup A$ is finite. $\mathscr{T}$ is called *finitely axiomatised* if A is finite.

Since each element of C is a variable of some axiom, and since each axiom involves only finitely many variables, we have either that C and A are both finite, or that $|C| \leqslant |A|$. If $\mathscr{R} \cup A$ is infinite, then $|\mathscr{L}(\mathscr{T})| = |\mathscr{R} \cup A|$, while $\mathscr{L}(\mathscr{T})$ is countable if $\mathscr{R} \cup A$ is finite. We remark that a relation symbol not occurring in any axiom would be of little interest, as it could be interpreted as any relation and so could occur in a theorem of $\mathscr{T}$ only in an essentially trivial way. It would not be a serious restriction to require every relation symbol to appear in some axiom, in which case we would have either $\mathscr{R}$ and A finite or $|\mathscr{R}| \leqslant |A| = |\mathscr{R} \cup A|$. When A is finite, the actual value of $|A|$ is of no real interest, because an axiom set $A = \{a_1, \ldots, a_n\}$ can always be replaced by $A' = \{a_1 \wedge \cdots \wedge a_n\}$ without making any essential change in the theory.

The following theorem is the main result of the present chapter, and is in fact the fundamental theorem of model theory.

Theorem 3.18. (Löwenheim-Skolem Theorem). *Let $\mathscr{T}$ be a first-order theory of cardinal χ, and let $\aleph$ be any infinite cardinal such that $\aleph \geqslant \chi$. Suppose $\mathscr{T}$ has an infinite model. Then $\mathscr{T}$ has a model of cardinal $\aleph$.*

Proof: Suppose $\mathscr{T} = (\mathscr{R}, A, C)$. Choose some set $V_0 \supset C$ such that $|V_0 - C| = \aleph$. Then $|P(V_0, \mathscr{R})| = \aleph$. Put

$$A'_0 = I \cup A \cup \{\sim\mathscr{I}(x, y) | x, y \in V_0 - C, x \neq y\}.$$

This gives a theory $\mathscr{T}' = (\mathscr{R}, A'_0, V_0)$ which we prove consistent. If $\mathscr{T}'$is inconsistent, then F is provable from A and some finite subset of $\{\sim\mathscr{I}(x, y)| x, y \in V_0 - C, x \neq y\}$, which contradicts the hypothesis that $\mathscr{T}$ has an infinite model. Therefore $\mathscr{T}'$ is consistent.

We follow the method used to prove the Satisfiability Theorem (cf Lemma 4.14 of Chapter IV), and construct inductively sets V_n, A'_n and A_n. We put

$$V_{n+1} = V_n \cup \{t_q^{(n)} | q(x) \in P(V_n, \mathscr{R}), (\exists x)q(x) \in A_n\},$$
$$A'_{n+1} = A_n \cup \{q(t_q^{(n)}) | q(x) \in P(V_n, \mathscr{R}), (\exists x)q(x) \in A_n\},$$

and take for A_{n+1} a maximal consistent subset of $P(V_{n+1}, \mathscr{R})$ containing A'_{n+1}. We put $V^* = \bigcup_n V_n$, $A^* = \bigcup_n A_n$, $P^* = P(V^*, \mathscr{R}) = \bigcup_n P(V_n, \mathscr{R})$.

Since $A_0 \supseteq A'_0$ is a maximal consistent subset of $P(V_0, \mathscr{R})$, and since $|P(V_0, \mathscr{R})| = \aleph$, we have $|A_0| = \aleph$. Then, from $|P(V_n, \mathscr{R})| = |A_n| = \aleph$, it follows that $|V_{n+1}| = \aleph$, and so that $|P(V_{n+1}, \mathscr{R})| = |A_{n+1}| = \aleph$. By induction, $|P(V_n, \mathscr{R})| = \aleph$ for all n, and therefore $|P^*| = \aleph$.

As in the proof of the Satisfiability Theorem, we construct an interpretation (P^*, φ, ψ, v) for which $v(A^*) = \{1\}$. In this interpretation, $\psi\mathscr{I}$ is an equivalence relation, and by replacing elements of P^* by their equivalence classes, we obtain a proper interpretation $(P^*/\psi\mathscr{I}, \bar{\varphi}, \bar{\psi}, v)$. Restricting $\bar{\varphi}$ to $\mathscr{L}(\mathscr{T}')$ gives a model M of $\mathscr{T}'$. Since $|P^*| = \aleph$, $|M| = |P^*/\psi\mathscr{I}| \leqslant \aleph$. But the construction of A'_0 ensures that any model of $\mathscr{T}'$ has cardinal at least $\aleph$. Therefore $|M| = \aleph$. Restricting to $\mathscr{L}(\mathscr{T})$ converts M into a model of $\mathscr{T}$. □

Corollary 3.19. *If the first-order theory $\mathscr{T}$ has an infinite model, then $\mathscr{T}$ is not categorical.* (Proof obvious.)

Corollary 3.20. *Suppose the theory $\mathscr{T}$ has cardinal χ, and has only infinite models. Suppose also that $\mathscr{T}$ is categorical in some infinite cardinal $\aleph \geqslant \chi$. Then $\mathscr{T}$ is complete.*

Proof: Let $p \in \mathscr{L}(\mathscr{T})$, and suppose that neither p nor $\sim p$ is a theorem of $\mathscr{T} = (\mathscr{R}, A, C)$. Since $\sim p$ is not a theorem, $\mathscr{T}$ has a model (infinite) in which p is true, and so the theory $\mathscr{T}' = (\mathscr{R}, A \cup \{p\}, C)$ has an infinite model. Since $|\mathscr{T}'| \leqslant \aleph$, $\mathscr{T}'$ has a model M' of cardinal $\aleph$. Similarly $\mathscr{T}'' = (\mathscr{R}, A \cup \{\sim p\}, C)$ has a model M'' of cardinal $\aleph$. But M' and M'' are each models of $\mathscr{T}$ of cardinal $\aleph$, and hence are isomorphic, contrary to p being true in M' and false in M''. □

Exercises

3.21. A dense linearly ordered set is a non-empty set with a binary relation $<$ such that

(a) for all x, y, exactly one of $x < y$, $x = y$, $y < x$ holds,
(b) if $x < y$ and $y < z$, then $x < z$,
(c) if $x < y$, then there exists z such that $x < z < y$,
(d) for each x, there exist y, z such that $y < x < z$.

Using a binary relation symbol ℓ, with $\ell(x, y)$ to be thought of as $x < y$, and also $\mathscr{I}$, construct a finite theory $\mathscr{D}$ whose models are precisely the dense linearly ordered sets. Show that every model of $\mathscr{D}$ is infinite. Prove that $\mathscr{D}$ is categorical in cardinal $\aleph_0$. (Hint: given two countable models of $\mathscr{D}$, enumerate each domain, and then define inductively a mapping, preserving $<$. Show that this map is onto by proving that there can be no first element in any omitted subset of the range space.) Deduce that $\mathscr{D}$ is a complete theory.

3.22. The theory $\mathscr{T} = (\mathscr{R}, A, C)$ has a finite model of cardinal χ. Show that there exists $p \in \mathscr{L}(\mathscr{T})$ such that the models of $\mathscr{T}' = (R, A \cup \{p\}, C)$ are precisely those models of $\mathscr{T}$ which have cardinal χ.

3.23. The theory $\mathscr{T} = (\mathscr{R}, A, C)$ has a model of infinite cardinal χ. Show that there is no subset T of $\mathscr{L}(\mathscr{T})$ such that $\mathscr{T}' = (\mathscr{R}, A \cup T, C)$ is a consistent theory, all of whose models have cardinal χ.

3.24. K is a field. Construct an elementary theory $\mathscr{V}_K$ of vector spaces over K, and, by adding extra axioms, an elementary theory $\mathscr{V}_K^\infty$ of infinite vector spaces over K. Show that $\mathscr{V}_K$ is categorical in any infinite cardinal greater than $|K|$. Hence show that $\mathscr{V}_K^\infty$ is complete. Show that $\mathscr{V}_K$ is not complete.

3.25. $\mathscr{T} = (\mathscr{R}, A, C)$ is a complete theory, which has a finite model $M = (M, \varphi, \psi)$ of cardinal n.

(i) Prove that every model of $\mathscr{T}$ has cardinal n.

Let $M' = (M', \varphi', \psi')$ be another model of $\mathscr{T}$, and let $\alpha : M \to M'$ be bijective. We say α preserves constants if $\alpha\varphi(c) = \varphi'(c)$ for all $c \in C$. We say that α preserves the relation $r \in \mathscr{R}_t$ if, for all $(m_1, \ldots, m_t) \in M^t$, we have $(m_1, \ldots, m_t) \in \psi r$ if and only if $(\alpha m_1, \ldots, \alpha m_t) \in \psi' r$. We say that α preserves the subset $\mathscr{S}$ of $\mathscr{R}$ if it preserves every $r \in \mathscr{S}$.

(ii) Show that a bijective map $\alpha: M \to M'$ is an isomorphism of models if and only if it preserves constants and preserves $\mathscr{R}$.

Let $a_1, \ldots, a_n$ be the elements of M, so numbered that $\varphi(C) = \{a_1, \ldots, a_k\}$, and let $c_1, \ldots, c_k \in C$ be such that $\varphi(c_i) = a_i$.

(iii) Show that a bijective map α preserves constants if and only if $\alpha(a_i) = \varphi'(c_i)$ for $i = 1, \ldots, k$.

For $r \in \mathscr{R}_t$ and $(i_1, \ldots, i_t) \in \mathbf{Z}_n^t$, we put

$$q(i_1, \ldots, i_t) = \begin{cases} r(x_{i_1}, \ldots, x_{i_t}) \text{ if } (a_{i_1}, \ldots, a_{i_t}) \in \psi r, \\ \sim r(x_{i_1}, \ldots, x_{i_t}) \text{ if } (a_{i_1}, \ldots, a_{i_t}) \notin \psi r, \end{cases}$$

and
$$r^*(x_1, \ldots, x_n) = \bigwedge_{(i_1, \ldots, i_t)} q(i_1, \ldots, i_t).$$

Write $\operatorname{dist}(x_1, \ldots, x_n)$ for $\sim\mathscr{I}(x_1, x_2) \wedge \sim\mathscr{I}(x_1, x_3) \wedge \cdots \wedge \sim\mathscr{I}(x_1, x_n) \wedge \cdots \wedge \sim\mathscr{I}(x_{n-1}, x_n)$.

(iv) Show that

$$(\exists x_{k+1}) \cdots (\exists x_n)(\operatorname{dist}(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n) \wedge r^*(c_1, \ldots, c_k, x_{k+1}, \ldots, x_n))$$

is a theorem of $\mathscr{T}$, and hence show that there exists $\alpha: M \to M'$ which is bijective, preserves constants and preserves r. Extend this argument to show for any finite subset $\mathscr{S}$ of $\mathscr{R}$, that there is a bijective map $\alpha: M \to M'$ preserving constants and $\mathscr{S}$.

(v) Using the fact that there are only finitely many bijective maps $\alpha: M \to M'$, and observing that if some given α preserving constants is not an isomorphism, then there is some $r \in \mathscr{R}$ not preserved by α, prove that M and M' are isomorphic. Hence prove that $\mathscr{T}$ is categorical.

§4 Reduction of Quantifiers

In any study of the decidability properties of a theory $\mathscr{T} = (\mathscr{R}, A, C)$, one expects those elements $q \in P(\mathscr{T})$ which involve no quantifiers to pose the least difficulty. If $q \in P(\mathscr{T})$ is quantifier-free, it is a propositional combination of primitive propositions $r(v_1, \ldots, v_n)$ $(r \in \mathscr{R}_n, v \in V)$, whose truth or falsity for any given interpretation of P is easily determined. Truth functions then decide the truth or falsity of q. There are theories $\mathscr{T}$ having the property that, for any $p \in P(\mathscr{T})$, a quantifier-free element $q \in P(\mathscr{T})$ can be found such that $\mathscr{T} \vdash p \Leftrightarrow q$ and $\operatorname{var}(q) \subseteq \operatorname{var}(p)$. Such a process of quantifier elimination could be useful in investigating the completeness or decidability of $\mathscr{T}$. In practice, it is rarely possible to eliminate quantifiers completely, and one must be content with a quantifier-reduction procedure. The resulting element q is then a propositional combination of relatively simple propositions which possibly involve quantifiers. We shall call these "simple" elements of $P(\mathscr{T})$ *primary propositions.*

Definition 4.1. Let $\Pi \subseteq P(\mathscr{T})$. We say that $\mathscr{T}$ *admits* Π*-reduction of quantifiers* if there is a process which assigns to each $p \in P(\mathscr{T})$ an element $q \in P(\mathscr{T})$ such that

(i) q is a propositional combination of elements of Π,

(ii) $\operatorname{var}(q) \subseteq \operatorname{var}(p)$,

and

(iii) $\mathscr{T} \vdash p \Leftrightarrow q$.

The utility of such a reduction procedure for any investigation depends on the relative simplicity of the elements of Π compared to the elements of $P(\mathscr{T})$. Every theory admits the useless reduction given by $\Pi = P(\mathscr{T})$ and $q = p$. We give a more helpful example.

Example 4.2. Let $\mathscr{E} = (\{\mathscr{I}\}, \varnothing, \varnothing)$ be the theory of equality, with $V = V_0 = \{x_n | n \in \mathrm{N}\}$. We write $\mathscr{I}(x, y)$ as $x = y$, and $\sim\mathscr{I}(x, y)$ as $x \neq y$. As the set of primary propositions, we take

$$\Pi = \{\mathrm{al}(n) | n \in \mathrm{N}\} \cup \{x_i = x_j | i, j \in \mathrm{N}\}.$$

We introduce an abbreviation which we shall use in describing the reduction process. For $1 \leqslant r \leqslant s$, put

$$\mathrm{dist}_r(x_1, \ldots, x_s) = \bigvee_{\alpha} \left(\bigwedge_{\imath < \jmath \leqslant r} (x_{\alpha_i} \neq x_{\alpha_j}) \wedge \bigwedge_{i=r+1}^{s} \left(\bigvee_{j=1}^{r} (x_{\alpha_i} = x_{\alpha_j}) \right) \right),$$

where α ranges over the permutations $(\alpha_1, \ldots, \alpha_s)$ of $(1, \ldots, s)$. Observe that $\mathrm{dist}_r(x_1, \ldots, x_s)$ is true in an interpretation if and only if the interpretation of $x_1, \ldots, x_s$ gives exactly r distinct elements of the model. Now put

$$\mathrm{only}(x_1, \ldots, x_s) = \bigvee_{r=1}^{s} (\mathrm{dist}_r(x_1, \ldots, x_s) \wedge \sim\mathrm{al}(r + 1)),$$

and observe that this is true for an interpretation if and only if the interpretations of $x_1, \ldots, x_s$ are all the elements of the model. (Thus, $\mathrm{only}(x_1, \ldots, x_s)$ is true if and only if $(\forall x)((x = x_1) \vee (x = x_2) \vee \cdots \vee (x = x_s))$ is true.) It is clear that $\mathrm{only}(x_1, \ldots, x_s)$ is a propositional combination of elements of Π.

The following set of instructions constitutes the reduction process:

Step 0. If p is quantifier-free, put $q = p$ and stop. Otherwise, express p in prenex normal form (see Exercise 4.18 of Chapter IV) $Q_1 Q_2 \cdots Q_r p_1$, where the Q_i are quantifiers and p_1 is quantifier-free (and hence a propositional combination of elements of the form $x_i = x_j$).

Step 1. We have $p = Q_1 Q_2 \cdots Q_r p_1$, where the Q_i are quantifiers and p_1 is a propositional combination of elements of Π. If $r = 0$, put $q = p$ and stop. If Q_r is a universal quantifier $(\forall x)$, proceed to Step 2. If Q_r is an existential quantifier $(\exists x)$, then replace Q_r by $\sim(\forall x)\sim$.

Step 2. We have $p = Q(\forall x)p_1$, where Q consists of a (possibly empty) string of quantifiers and possibly a negation, and p_1 is a propositional combination of elements of Π. If $x \notin \operatorname{var}(p_1)$, replace p by Qp_1 and begin again at Step 1. If $x \in \operatorname{var}(p_1)$, express p_1 in conjunctive normal form (see Exercise 3.10 of Chapter III):

$$p_1 = a_1 \wedge a_2 \wedge \cdots \wedge a_k,$$

where $a_j = d_{1j} \vee d_{2j} \vee \cdots \vee d_{n_j j}$, with each d_{ij} either a primary proposition or the negation of a primary proposition.

Step 3. We have $p = Q(\forall x)p_1$, with $p_1 = \bigwedge_{j=1}^{k} \bigvee_{i=1}^{n_j} d_{ij}$, where each d_{ij} is primary or the negation of a primary proposition. For each $j = 1, 2, \ldots, k$, delete a_j from p_1 if there is an i such that $d_{ij} = (v = v)$ for some $v \in V$, unless this holds for all j, in which case replace $(\forall x)p_1$ by $F \Rightarrow F$ and begin again at Step 1.

Step 4. We have $p = Q(\forall x)p_1$, with $p_1 = \bigwedge_{j=1}^{k} \bigvee_{i=1}^{n_j} d_{ij}$, where d_{ij} is primary or the negation of a primary proposition. For each $j = 1, 2, \ldots, k$, delete from a_j every d_{ij} of the form $(v \neq v)$ with $v \in V$, unless for some j every d_{ij} has this form, in which case replace $(\forall x)p_1$ by F and begin again at Step 1.

Step 5. We have $p = Q(\forall x)p_1$, with $p_1 = \bigwedge_{j=1}^{k} \bigvee_{i=1}^{n_j} d_{ij}$, where d_{ij} is primary or the negation of a primary proposition, and where no d_{ij} has the form $(v = v)$ or the form $(v \neq v)$. Put $a_j' = \bigvee\{d_{ij} \mid x \in \operatorname{var}(d_{ij})\}$, and $a_j'' = \bigvee\{d_{ij} \mid x \notin \operatorname{var}(d_{ij})\}$, so that $a_j = a_j' \vee a_j''$. Since the only elements π of Π for which $x \in \operatorname{var}(\pi)$ are the elements $x = v$ for $v \in V$, it follows that each non-empty a_j' has the form

$$a_j' = (x = v_1) \vee (x = v_2) \vee \cdots \vee (x = v_s) \vee (x \neq w_1) \vee \cdots \vee (x \neq w_t)$$

for elements $v_1, \ldots, v_s, w_1, \ldots, w_t$ of $V - \{x\}$. (Terms $x = x$ or $x \neq x$ are excluded by Steps 3, 4.) For each j such that a_j' is non-empty, then

(a) if $t = 0$, replace a_j' by only $(v_1, \ldots, v_s)$.
(b) if $t = 1$ and $s = 0$, replace a_j' by F.
(c) if $t > 0$ and $(s, t) \neq (0, 1)$, replace a_j' by
$$(w_1 \neq w_2) \vee (w_1 \neq w_3) \vee \cdots \vee (w_1 \neq w_t) \vee (v_1 = w_1) \vee \cdots \vee (v_s = w_1).$$

Finally, delete $(\forall x)$. Now return to Step 1.

In the above procedure, each step replaces the given proposition by one equivalent to it (i.e., true for precisely the same interpretations). In Step 5, for example, we note that $(\forall x)(a_1 \wedge \cdots \wedge a_k)$ is equivalent to $(\forall x)a_1 \wedge (\forall x)a_2 \wedge \cdots \wedge (\forall x)a_k$, and consider each $(\forall x)a_j$ separately. At each return to Step 1, the number of quantifiers in the prefix has been reduced, so the process must stop.

We illustrate the use of quantifier reduction by proving that $\mathscr{E}$ is decidable.

Theorem 4.3. *The theory* $\mathscr{E} = (\{\mathscr{I}\}, \varnothing, \varnothing)$ *is decidable.*

Proof. Let $p \in \mathscr{L}(\mathscr{E})$. The reduction process described above gives an element q, equivalent to p, which is a propositional combination of elements of Π such that $\mathrm{var}(q) \subseteq \mathrm{var}(p)$. Since $\mathrm{var}(p) = \varnothing$, q is a propositional combination of elements of the form $p_n = \mathrm{al}(n + 1)$ $(n \geqslant 1)$. Hence q is a propositional combination of $p_1, p_2, \ldots, p_k$ for some k. Let $f\colon \mathbf{Z}_2^k \to \mathbf{Z}_2$ be the corresponding truth function. Then $\mathscr{E} \vdash q$ if and only if $f(x_1, \ldots, x_k) = 1$ for all $(x_1, \ldots, x_k) \in \mathbf{Z}_2^k$ such that, for some n $(0 \leqslant n \leqslant k)$, $x_1 = x_2 = \cdots = x_n = 1$ and $x_{n+1} = x_{n+2} = \cdots = x_k = 0$. This is so because these are the only possible combinations of truth values for $p_1, \ldots, p_k$ in models of $\mathscr{E}$. □

We note that there is no need for a formal definition of decidability of a first-order theory when one is proving constructively that a particular theory is decidable—the proof is self-sufficient. Formality is required if one is to show the nonexistence of a decision process, as we shall do in Chapter IX. We also remark that the above result, on the decidability of the theory of equality, is not in conflict with the theorem of Kalmar mentioned in Chapter IV and proved in Chapter IX. Although the theory of equality involves a binary predicate symbol, it also includes the axioms of identity.

Exercise 4.4. Show that the theory $\mathscr{D}$ (Exercise 3.21) of dense linear order admits Π-reduction of quantifiers with $\Pi = \{(x_i = x_j), (x_i < x_j) \mid i, j \in \mathbf{N}\}$. Hence show that $\mathscr{D}$ is decidable and complete.

Chapter VI

Zermelo-Fraenkel Set Theory

§1 Introduction

All the ordinary mathematical systems are constructed in terms of sets. If we wish to study the reasoning used in mathematics, our model of mathematics must include some form of set theory, for otherwise our study must be restrictive. For example, Elementary Group Theory formalises almost nothing of group theory. The pervasive role of set theory in mathematics implies that any reasonable model of set theory will in effect contain a model of all of mathematics (including the mathematics of this book).

The informal way in which properties of sets are used in mathematics often means that one is aware of some of the more useful axioms of set theory without necessarily having seen or studied sets as an axiomatic theory. In those parts of mathematics where a careful account of set theory is needed, the axiomatisation usually chosen is the one known as Zermelo-Fraenkel Set Theory. We shall set out the axioms of this theory (which we denote by ZF) with some brief comments on the significance of the various axioms. We shall then see how this theory ZF may be formalised within $\mathrm{Pred}_{\mathscr{I}}(V, \mathscr{R})$. Finally, we shall consider the significance of some of the results of Chapter V for our formalised set theory. The reader interested in a more detailed account of ZF is referred to [4].

§2 The Axioms of ZF

ZF is the study of a single type of object. Objects of this type will be called sets. We shall admit another type of object, called a property of a set, but the objects which make up any set will themselves be sets. Since one customarily forms sets whose members are mathematical or physical objects of diverse types, the requirement that members of sets must themselves be sets may therefore seem restrictive. Experience has shown that with some exceptions (which can be accommodated by an extension of the theory), all the objects used in mathematics can be constructed as sets, while we can avoid the need to form sets of physical objects by assigning mathematical names to the objects and using the set of names.

In ZF, we study a single relationship[1] between sets. This relationship is called membership and will be denoted by $\in$. Thus $x \in y$ is read "(the set) x

[1] We cannot formalise this relationship as a set of pairs, for we are after all just beginning to define our set theory. Later, when we have constructed ZF, we shall see that the collection of pairs involved cannot be a set within ZF.

is a member of (the set) y", or "x belongs to y". We also study property relationships, which are of the form "the set x has the property π".

In the list of axioms of ZF which follows, some are described as axioms, others as axiom schemas. The distinction will be explained when we construct First-Order ZF.

(ZF1) Axiom of Extension. *If a and b are sets, and if for all sets x we have $x \in a$ if and only if $x \in b$, then $a = b$.*

Thus two sets are equal if and only if they have the same members. We shall write $a \subseteq b$ if $x \in a$ implies $x \in b$.

(ZF2) Axiom Schema of Subsets. *For any set a and any property π, there is a set b such that $x \in b$ if and only if $x \in a$ and has the property π.*

By (ZF1), this set is unique. We denote it by $\{x \in a | x \text{ has } \pi\}$. Assuming that at least one set a exists, we can form the set $\emptyset = \{x \in a | x \neq x\}$. Then for all x we have $x \notin \emptyset$. This set $\emptyset$, which is called the empty set, is independent of the choice of the set a used in its construction. By (ZF1), $\{x \in a_1 | x \neq x\} = \{x \in a_2 | x \neq x\}$. It is clear that for all sets b, $\emptyset \subseteq b$.

(ZF2) restricts the way in which a property may be used to form a set, and thereby, the Russell paradox is avoided. It used to be assumed that, for any property π, one could form the set of all objects with that property. Russell considered the property of not being a member of itself. If b is the set of all sets which are not members of themselves, then consideration of whether or not b is a member of itself leads at once to a contradiction. Using (ZF2), one can only form $b = \{x \in a | x \notin x\}$ starting from some given set a. We then find that $b \in b$ is impossible, hence $b \notin b$ and so $b \notin a$. The argument does not lead to a contradiction, but instead proves that for any a, there is a b such that $b \notin a$. Thus there is no set of all sets.

(ZF3) Axiom of Pairing. *If a and b are sets, then there exists a set c such that $a \in c$ and $b \in c$.*

Using (ZF2) with this set c, we can form the set $\{x \in c | x = a \text{ or } x = b\}$. This is independent of the particular set c having a and b as members, and we call $\{x \in c | x = a \text{ or } x = b\}$ the unordered pair whose members are a and b, and denote it by $\{a, b\}$. In the special case where $a = b$, (ZF2) asserts the existence of a set having a as a member. The unordered pair $\{a, a\}$ has only the one member a, and we denote it by $\{a\}$. The ordered pair (a, b) is now defined to be $\{\{a\}, \{a, b\}\}$.

Exercise 2.1. If $(a, b) = (c, d)$, prove $a = c$ and $b = d$. Make sure that your proof allows for the possibility that $a = b$.

For any two sets a, b, we can form $a \cap b = \{x \in a | x \in b\}$. For any non-empty set c, we can form $\cap c = \{x \in b | x \in a \text{ for all } a \in c\}$, where b is some member of c. $\cap c$ is, of course, independent of the choice of b.

Exercise 2.2. Prove that $a \cap b = b \cap a = \cap\{a, b\}$.

Although the axioms already given allow the formation of intersections, the formation of unions requires a further axiom.

(ZF4) Axiom of Union. *For every set c, there exists a set a such that, if $x \in b$ and $b \in c$, then $x \in a$.*

We can now form $\cup c = \{x \in a | x \in b \text{ for some } b \in c\}$ where a is as in (ZF4). $\cup c$ is again independent of the particular a used, so we write simply $\cup c = \{x | x \in b \text{ for some } b \in c\}$. For any sets a and b, we can form $a \cup b = \cup\{a, b\}$.

Exercise 2.3. Show that the ordered pairs (a, b) for which $a \in b$ do not form a set. (Assume that there is a set $e = \{(a, b) | a \in b\}$ and show that $\cup(\cup e)$ is the set of all sets.)

The formation of ordered pairs is permitted by the axioms so far given, but not the formation of the set of all ordered pairs of members of given sets. The next axiom remedies this deficiency.

(ZF5) Axiom of the Power Set. *For each set a, there exists a set b such that, if $x \subseteq a$, then $x \in b$.*

Using (ZF2), we obtain the existence of the power set of a: $\text{Pow}(a) = \{x \in b | x \subseteq a\} = \{x | x \subseteq a\}$, which is clearly independent of the choice of b.

(ZF5) allows formation of the cartesian product $a \times b = \{(x, y) | x \in a \text{ and } y \in b\}$. To show this, we need only produce a set c whose members include all the required ordered pairs (x, y). But $(x, y) = \{\{x\}, \{x, y\}\}$, $\{x\} \subseteq a \cup b$, $\{x, y\} \subseteq a \cup b$, and so both $\{x\}$ and $\{x, y\}$ are members of $\text{Pow}(a \cup b)$. Thus $\{\{x\}, \{x, y\}\} \subseteq \text{Pow}(a \cup b)$, and consequently $(x, y) \in \text{Pow}(\text{Pow}(a \cup b))$ for all $x \in a$ and $y \in b$.

With the cartesian product available, we can now define a relation between two sets a, b as a subset of $a \times b$, and then a function $f: a \to b$ as a special type of relation. The set of all functions from a to b can be constructed as a subset of $\text{Pow}(\text{Pow}(a \times b))$. For a set c, we define the cartesian product (of the members) of c by $\prod c = \{f: c \to \cup c | f(x) \in x \text{ for all } x \in c\}$.

Exercises

2.4. What is $\prod \varnothing$?

2.5. For any set a, prove that there is no surjective function $f: a \to \text{Pow}(a)$. (Consider $b = \{x \in a | x \notin f(x)\}$.)

Definition 2.6. The *successor* of the set x is the set $x^+ = x \cup \{x\}$. The set a is called a *successor set* if $\varnothing \in a$ and $x^+ \in a$ for all $x \in a$.

(ZF6) Axiom of Infinity. *There exists a successor set.*

This is the first axiom asserting unconditionally that sets exist. In particular, it asserts the existence of $\varnothing$ as this is used in the definition of a successor set. We can now define the set ω of natural numbers:

$$\omega = \{x | x \in a \text{ for every successor set } a\},$$

using (ZF2) and some successor set. We use the usual symbols

$$\begin{aligned}0 &= \varnothing,\\ 1 &= 0^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\} = \{0\},\\ 2 &= 1^+ = 1 \cup \{1\} = \{0, 1\},\\ 3 &= 2^+ = \{0, 1\} \cup \{2\} = \{0, 1, 2\},\end{aligned}$$

and so on. The set ω together with the usual operations of addition and multiplication will be denoted by **N**.

Exercises

2.7. The set n is called transitive if $x \in y$ and $y \in n$ imply that $x \in n$. Show that if n is transitive, then so is n^+.

2.8. If s is a successor set, show that $\{n \in s | n \text{ is transitive}\}$ is a successor set. For all $n \in \omega$, prove that n is transitive.

2.9. Given that $n = \{x \in \omega | x \subset n\}$ and that $n \in \omega$, show that $n^+ = \{x \in \omega | x \subset n^+\}$. Hence prove for all $n \in \omega$ that

(a) $n = \{x \in \omega | x \subset n\}$,
(b) $n \notin n$,
(c) for all $x \in n$, $n \nsubseteq x$.

($a \subset b$ means $a \subseteq b$ and $a \neq b$.)

2.10. Show that 0, 1, 2, ... are all different.

(ZF7) Axiom of Choice. *For each set a, there exists a function $f: \{x \in \text{Pow}(a) | x \neq \varnothing\} \to a$, such that for every non-empty subset x of a, $f(x) \in x$.*

The function f, called a *choice function*, selects from each non-empty subset of a, a member of that subset.

(ZF8) Axiom Schema of Replacement. *If π is a property of pairs of sets such that for all $x \in a$, (x, y) and (x, z) both having π implies that $y = z$, then there exists a set b such that $y \in b$ if and only if there is an $x \in a$ such that (x, y) has π.*

Intuitively, the property π defines a function on some subset of a, and b is the set of images under this function. But a function $f: a \to b$ is a subset of $a \times b$, and this requires b to be a set. The point of this axiom is that although we are not given a function in the formal sense, the type of correspondence it considers does in fact define a function.

(ZF9) Axiom Schema of Restriction. *If π is any property of sets and if there exists a set with π, then there exists a set a with π such that for all $x \in a$, x does not have π.*

(ZF9) excludes the possibility of an infinite sequence $a_1, a_2, \ldots$ of sets such that $a_{i+1} \in a_i$ for all i. To see this, simply take π to be the property of being the first member of some such sequence. By (ZF9), if there exists a set with this property π, then there exists a set a with π such that no member of a

has π. But then a is the first member a_1 of some such sequence $a_1 \ni a_2 \ni \cdots$, and clearly $a_2 \in a$ and has π. Thus there can be no sets with this property.

Exercises

2.11. Show that (ZF9) implies the Axiom of Regularity: For any set $a \neq \varnothing$, there exists $b \in a$ such that $b \cap a = \varnothing$.

2.12. From the Axiom of Regularity, prove that for every set a, $a \notin a$.

2.13. Prove that if $a \subseteq a \times a$, then $a = \varnothing$.

§3 First-Order ZF

We formalise ZF as a first-order theory, which we shall denote by $\mathscr{S}$. We take as relation symbols just $\mathscr{I}$, $\in$, both binary. We shall use no individual constants in our construction. Where axioms are obvious formalisations of the corresponding informal axioms, we set them down without comment. For ease of understanding, we shall write $x \in y$ and $x = y$ rather than the formally correct $\in(x, y)$ and $\mathscr{I}(x, y)$, and the negations of these statements will be written $x \notin y$ and $x \neq y$.

(ZF1) $(\forall a)(\forall b)(((\forall x)(x \in a \Leftrightarrow x \in b)) \Rightarrow a = b)$.

In the informal version of (ZF2), we used a property π of sets. The informal statement "x has property π." becomes for us the predicate $\pi(x)$, where π is an element of $P(V, \mathscr{R})$, the notation $\pi(x)$ simply describing the dependence of π upon x. (The notation $\pi(x)$ does not imply that $\operatorname{var}(\pi(x)) = \{x\}$.) For given $\pi(x)$, (ZF2) becomes

$$(\forall a)(\exists b)(\forall x)(x \in b \Leftrightarrow (x \in a \wedge \pi(x))),$$

but we must clearly restrict this by requiring that $b \notin \operatorname{var}(\pi(x))$. Moreover, the theory $\mathscr{S}$ is to be without constants, and $\operatorname{var}(\pi(x))$ could have members other than a and x. Thus, if $x_1, \ldots, x_r$ are these other variables, we take as our axiom

$$(\forall x_1) \cdots (\forall x_r)(\forall a)(\exists b)(\forall x)(x \in b \Leftrightarrow (x \in a \wedge \pi(x))).$$

To simplify the notation, we introduce the convention that if $p \in P(V, \mathscr{R})$ and $\operatorname{var}(p) = \{x_1, \ldots, x_n\}$, then $(\forall)p$ denotes $(\forall x_1) \cdots (\forall x_n)p$. The order in which $x_1, \ldots, x_n$ are taken will not matter in any use we make of this notation. Using this convention, the axiom schema becomes

(ZF2) $(\forall)(\forall a)(\exists b)(\forall x)(x \in b \Leftrightarrow (x \in a \wedge \pi(x)))$ *for all* $\pi(x) \in P(V, \mathscr{R})$ *such that* $b \notin \operatorname{var}(\pi(x))$.

Unlike (ZF1), which was a single element of $P(V, \mathscr{R})$, (ZF2) is an infinite collection of axioms, one for each $\pi(x) \in P(V, \mathscr{R})$ satisfying $b \notin \operatorname{var}(\pi(x))$. This is the reason for calling (ZF2) an axiom schema.

Exercise 3.1. Some later axioms of $\mathscr{S}$ will have the form $(\forall)(\exists a)(\forall x)(p(x) \Rightarrow x \in a)$ for certain elements $p(x) \in P(V, \mathscr{R})$. Show that $\{(\forall)(\exists a)(\forall x)(p(x) \Rightarrow x \in a)\} \vdash_{\mathscr{S}} (\forall)(\exists a)(\forall x)(p(x) \Leftrightarrow x \in a)$.

We introduce further useful abbreviations. We write $a \subseteq b$ for $(\forall x)(x \in a \Rightarrow x \in b)$, $a = \{x|p(x)\}$ for $(\forall x)(x \in a \Leftrightarrow p(x))$, $a = \{a_1, \ldots, a_n\}$ for $(\forall x)(x \in a \Leftrightarrow (x = a_1 \vee \cdots \vee x = a_n))$, and $c = (a, b)$ for $c = \{\{a\}, \{a, b\}\}$, which itself is an abbreviation whose meaning has been explained. In particular, $a = \varnothing$ is an abbreviation for $(\forall x)(x \notin a)$. We may now write down relatively concise formal versions of four more axioms.

(ZF3) $(\forall a)(\forall b)(\exists c)(a \in c \wedge b \in c)$.
(ZF4) $(\forall c)(\exists a)(\forall x)((\exists b)(x \in b \wedge b \in c)) \Rightarrow x \in a)$.
(ZF5) $(\forall a)(\exists b)(\forall x)(x \subseteq a \Rightarrow x \in b)$.
(ZF6) $(\exists a)((\exists b)(b = \varnothing \wedge b \in a)) \wedge (\forall x)(x \in a \Rightarrow (\exists y)(y = x \cup \{x\} \wedge y \in a)))$.

Exercises

3.2. Prove $\mathscr{S} \vdash (\forall a)(\forall b)(\exists c)(c = \{a, b\})$.

3.3. Formalise and prove the formal result that if $(a, b) = (c, d)$, then $a = c$ and $b = d$.

3.4. Prove $\mathscr{S} \vdash c \neq \varnothing \Rightarrow (\exists d)(\forall x)(x \in d \Leftrightarrow (\forall y)(y \in c \Rightarrow x \in y))$.

In (ZF6), $y = x \cup \{x\}$ is of course an abbreviation for $(\forall z)(z \in y \Leftrightarrow (z = x \vee z \in x))$. We further preserve our informal notations for certain sets by writing $b = \text{Pow}(a)$ for $(\forall x)(x \in b \Leftrightarrow x \subseteq a)$ and $c = a \times b$ for $(\forall x)(x \in c \Leftrightarrow (\exists y)(\exists z)(x = (y, z) \wedge y \in a \wedge z \in b))$.

To make possible a formal version of (ZF7) of reasonable length, we introduce three more abbreviations. We shall write $(\exists! x)p(x)$ for$(\exists x)(p(x) \wedge (\forall y)(p(y) \Leftrightarrow y = x))$ (as in Chapter V), $f\colon a \to b$ for

$$((\exists c)((c = a \times b) \wedge (f \subseteq c))) \wedge (\forall x)(x \in a \Rightarrow (\exists! y)((y \in b) \wedge (\exists z)(z = (x, y) \wedge z \in f))),$$

and $y = f(x)$ for $(\exists z)(z = (x, y) \wedge z \in f)$.

(ZF7) $(\forall a)(\forall b)((b = \{x|(x \subseteq a) \wedge (x \neq \varnothing)\}) \Rightarrow (\exists f)((f\colon b \to a) \wedge (\forall y)(\forall z)(z = f(y) \Rightarrow z \in y)))$.

(ZF8) *For every* $p(x, y) \in P(V, \mathscr{R})$, $(\forall)(((\forall x)(\forall y)(\forall z)((x \in a \wedge p(x, y) \wedge p(x, z)) \Rightarrow y = z)) \Rightarrow (\exists b)(\forall y)(y \in b \Leftrightarrow (\exists x)(x \in a \wedge p(x, y))))$.

(ZF9) *For every* $p(x) \in P(V, \mathscr{R})$,

$$(\forall)((\exists x)p(x) \Rightarrow (\exists a)(p(a) \wedge (\forall x)(x \in a \Rightarrow \sim p(x)))).$$

This completes the formalisation of the axioms of our informal set theory, and so completes the list of mathematical axioms of our first-order theory $\mathscr{S}$.

By its construction, $\mathscr{S}$ is clearly a consistent theory if our informal set theory is consistent, because any proof of F in $\mathscr{S}$ has an informal equivalent. Since this book (and also much of mathematics) is written in the context of the informal set theory of ordinary mathematics, and since all of this is destroyed if that set theory is inconsistent, we assume the consistency of informal set theory. With this assumption, $\mathscr{S}$ is a consistent theory.

We now observe that the language $\mathscr{L}(\mathscr{S})$ is in fact independent of the choice of the infinite set V of variables used, for if V_0 is a countable subset of V, and if $p \in P(V, \mathscr{R})$ has $\operatorname{var}(p) = \varnothing$, then $p \in P(V_0, \mathscr{R})$, and the result follows on recalling that there are no individual constants in our construction of $\mathscr{S}$. We may therefore suppose that V is countable. Since $\mathscr{R} = \{\mathscr{I}, \in\}$, it follows that $P(V, \mathscr{R})$ is countable and hence that $\mathscr{S}$ is countable. By the Löwenheim-Skolem Theorem, $\mathscr{S}$ has a countable model.

A theorem of ordinary set theory asserts the existence of uncountable sets, and this theorem (with its proof) can be formalised in the theory $\mathscr{S}$. Hence there exists a countable model of a theory which has as a theorem the existence of uncountable sets! The paradox is resolved when we realise that it arises by using the word "set" in two ways. Let us distinguish words used in their ordinary sense from the same words used in the sense of the model by using the adjectives real or model respectively. "$\mathscr{S}$ has a countable model" then becomes "$\mathscr{S}$ has a real countable model", i.e., there is a real function from the real set of natural numbers onto the underlying set of the model. For this model, every model set is at most real countable. But a model set is model countable only if there is a model function from the model set of natural numbers onto it, and the real function which counts it need not be a model function.

§4 The Peano Axioms

We have seen how the natural numbers may be constructed in terms of set theory. We now give an axiomisation of the natural numbers, and study the relationship between this axiomatic system and Zermelo-Fraenkel set theory.

Since addition and multiplication can be defined in terms of the successor function[2], it is sufficient to axiomatise this function. We denote the successor of x by $s(x)$. The Peano axioms for the natural number system $\mathbf{N}$ are:

P_1: *0 is a natural number.*

P_2: *If x is a natural number, then $s(x)$ is a uniquely determined natural number* (i.e., s is a function $s\colon \mathbf{N} \to \mathbf{N}$).

P_3: *If x, y are natural numbers and if $s(x) = s(y)$, then $x = y$.*

P_4: *For each natural number x, $s(x) \neq 0$.*

P_5: *If π is any property such that 0 has π, and such that if x has π then $s(x)$ has π, then every natural number has π.*

[2] However, addition and multiplication are not definable within the theory $\mathscr{P}$ we are about to construct. To be able to formalise their definitions, we have to add to $\mathscr{R}$ relation symbols for addition and multiplication. See Exercises 4.2–4.10.

It is a well-known theorem that these axioms determine the system N to isomorphism, i.e., if sets A, A', with functions s, s' respectively, each satisfy the axioms, then there exists a bijective function $f: A \to A'$ such that $f(0) = 0'$ and $f(s(a)) = s'(f(a))$ for all $a \in A$. An informal proof of this runs as follows. We define $f(0) = 0'$, and, if $f(a) = a'$, define $f(s(a)) = s'(a')$. Taking $\pi(a)$ to be the property that $f(a)$ is uniquely defined by this rule, P_5 then gives the result that f is a function from A to A'. Similarly, we obtain a function $g: A' \to A$. Taking now $\pi(a)$ to be $g(f(a)) = a$, P_5 gives the result that gf is the identity. Similarly fg is the identity, and so f is the required isomorphism.

The Peano axioms are easily formalised as a first-order theory $\mathscr{P}$. We take one unary relation symbol θ, with $\theta(x)$ to mean $x = 0$, and one binary relation symbol s, with $s(x, y)$ to mean x is the successor of y. The axioms then become

P_1: $(\exists! x)\theta(x)$.
P_2: $(\forall x)(\exists! y)s(y, x)$.
P_3: $(\forall x)(\forall y)(\forall z)((s(z, x) \wedge s(z, y)) \Rightarrow x = y)$.
P_4: $(\forall x)(\forall y)(s(x, y) \Rightarrow \sim\theta(x))$.
P_5: $(\forall)(((\exists x)(\theta(x) \wedge \pi(x)) \wedge (\forall y)(\forall z)(\pi(z) \wedge s(y, z) \Rightarrow \pi(y))) \Rightarrow (\forall y)\pi(y))$, *for all* $\pi(x) \in P(V, \mathscr{R})$ *such that* $y, z \notin \operatorname{var}(\pi(x))$.

$\mathscr{P}$ is clearly a countable theory, and has N as a model. By the Löwenheim-Skolem Theorem, $\mathscr{P}$ is not categorical. This result appears to contradict the theorem that the Peano axioms determine N to isomorphism. But in formalising P_5, we have restricted the application of the axiom to those properties π which are expressible in terms of s and θ, and the properties π used in the uniqueness proof are certainly not of this form. This argument is however only part of the whole story.

Within $\mathscr{P}$, we cannot hope to formalise a proof of the uniqueness theorem. We cannot even state the theorem in $\mathscr{L}(\mathscr{P})$. We need set theory for this, so let us reformulate the Peano axioms within our formal set theory $\mathscr{S}$, as a set of assumptions on a triple $(N, s, 0)$ of sets. We shall take $\mathscr{P}(N, s, 0)$ to be the subset of the first-order algebra of $\mathscr{S}$ consisting of the elements

P_1: $0 \in N$,
P_2: $s: N \to N$,
P_3: $(\forall x)(\forall y)(\forall z)((z = s(x) \wedge z = s(y)) \Rightarrow x = y)$,
P_4: $(\forall x) \sim (0 = s(x))$,
and all elements of the form
P_5: $(\forall)((\pi(0) \wedge (\forall x)(\forall y)((y = s(x) \wedge \pi(x)) \Rightarrow \pi(y))) \Rightarrow (\forall z)\pi(z))$,
where $y, z \notin \operatorname{var}(\pi(x))$.

We write $(N, s, 0) \simeq (N', s', 0')$ as an abbreviation for

$$(\exists f)((f: N \to N') \wedge (\forall x)(\forall y)(\forall z)((z = f(x) \wedge z = f(y)) \Rightarrow x = y)$$
$$\wedge (\forall x)(x \in N' \Rightarrow (\exists y)(x = f(y))) \wedge (\forall x)(\forall y)(\forall z)(\forall t)((y = s(x)$$
$$\wedge z = f(x) \wedge t = f(y)) \Rightarrow t = s(z))).$$

It can be shown that $\mathscr{P}(N, s, 0) \cup \mathscr{P}(N', s', 0') \vdash_{\mathscr{S}} (N, s, 0) \simeq (N', s', 0')$.

Hence within $\mathscr{S}$, the Peano axioms as now formulated in fact determine N to isomorphism. The axioms of $\mathscr{S}$, together with the assumptions $\mathscr{P}(N, s, 0)$, still do not determine N to isomorphism in the sense of our metalogic. There are non-isomorphic models of $\mathscr{S}$, and the systems of natural numbers within these models may well be non-isomorphic. Our theorem asserts that models of the natural numbers within a given model of $\mathscr{S}$ are isomorphic. Our informal proof worked because we were working within an assumed set theory.

Exercises

4.1. Rephrase our very informal proof of the uniqueness of the natural numbers more carefully in terms of informal axiomatic set theory. (This may be found in [12].) Note that the function f to be constructed is a subset of $N \times N'$ and must be constructed in a way permitted by the axioms. (The inductive construction of f needs justification.) Set out the steps of the argument in sufficient detail for it to become clear that it can be formalised to give a proof that $\mathscr{P}(N, s, 0) \cup \mathscr{P}(N', s', 0') \vdash_{\mathscr{S}} (N, s, 0) \simeq (N', s', 0')$.

4.2. Addition is usually defined in terms of the successor function by

(i) $x + 0 = x$, and
(ii) $x + s(y) = s(x + y)$.

Assuming the informal Peano axioms, show that (i) and (ii) define a function $+ : \mathrm{N} \times \mathrm{N} \to \mathrm{N}$, and that

(a) $0 + x = x$,
(b) $s(x) + y = s(x + y)$,
(c) $x + y = y + x$,
(d) $(x + y) + z = x + (y + z)$,
(e) $x + y = x + z$ implies $y = z$.

Give a similar definition of multiplication in terms of addition and the successor function, and establish its basic properties.

In the following exercises, $x_i = n$ (where $n \in \mathrm{N}$) is used as an abbreviation for

$$(\exists y_0)(\exists y_1) \cdots (\exists y_{n-1})(\theta(y_0) \wedge s(y_1, y_0) \wedge \cdots \wedge s(x_i, y_{n-1})$$

if $n > 0$, and means $\theta(x_i)$ if $n = 0$. The expression $x_i = x_j + n$ means $x_i = x_j$ if $n = 0$, $s(x_i, x_j)$ if $n = 1$, and

$$(\exists y_1)(\exists y_2) \cdots (\exists y_{n-1})(s(y_1, x_j) \wedge s(y_2, y_1) \wedge \cdots \wedge s(x_i, y_{n-1}))$$

if $n > 1$.

4.3. $\mathscr{P}^*$ is the theory formed from $\mathscr{P}$ by replacing the induction axiom scheme P_5 by

$$P^*_{5,0}: (\forall x)(((\forall y) \sim s(x, y)) \Rightarrow \theta(x)),$$
$$P^*_{5,n}: (\forall x)(x \neq x + n) \qquad (n > 0).$$

$M_I = \mathrm{N} \cup (\mathbf{Z} \times I)$, where I is some index set. For $m \in M_I$, $\theta(m)$ is interpreted as true if and only if $m = 0 \in \mathrm{N}$, and $s(m_1, m_2)$ is true if and only if either $m_2 \in \mathrm{N}$ and $m_1 = m_2 + 1$, or $m_2 = (z, i)$, where $z \in \mathbf{Z}$ and $i \in I$, and $m_1 = (z + 1, i)$. Show that M_I is a model of $\mathscr{P}^*$, and that every model of $\mathscr{P}^*$ is isomorphic to M_I for some I.

4.4. Prove that every theorem of $\mathscr{P}^*$ is a theorem of $\mathscr{P}$. Hence show that every model of $\mathscr{P}$ is isomorphic to M_I for some I.

4.5. Show that $\mathscr{P}^*$ admits Π-reduction of quantifiers, where

$$\Pi = \{x_i = x_j + n, x_i = n | i, j, n \in \mathrm{N}\}.$$

Hence prove that $\mathscr{P}^*$ is decidable and complete.

4.6. Let $\pi(x_0, x_1, \dots, x_n) \in P(\mathscr{P}) = P(\mathscr{P}^*)$, and let $a_1, \dots, a_n \in M_I$. Put

$$X = \{m \in M_I | \pi(m, a_1, \dots, a_n) \text{ is true in } M_I\}.$$

Using the Π-reduction of quantifiers, show that X is either finite or has finite complement in M_I. Hence prove that M_I satisfies the induction axiom scheme P_5, and so is a model of $\mathscr{P}$.

4.7. From the completeness of $\mathscr{P}^*$ and the fact that every theorem of $\mathscr{P}^*$ is a theorem of $\mathscr{P}$, deduce that every theorem of $\mathscr{P}$ is a theorem of $\mathscr{P}^*$. Hence prove that every M_I is a model of $\mathscr{P}$.

4.8. (Proof that M_I is a model of $\mathscr{P}$ not using reduction of quantifiers.) Show that $\mathscr{P}$ and $\mathscr{P}^*$ are α-categorical for every uncountable cardinal α, and so are complete. As in 4.7, deduce that every M_I is a model of $\mathscr{P}$.

4.9. The theory $\mathscr{A}$ consists of $\mathscr{P}$ together with a ternary relation symbol a and the additional axioms

$$(\forall x)(\forall y)(\exists! z)a(x, y, z),$$
$$(\forall x)(\forall y)(\theta(y) \Rightarrow a(x, y, x)),$$
$$(\forall x)(\forall y)(\forall z)(\forall t)(\forall u)(s(z, y) \wedge a(x, y, t) \wedge a(x, z, u) \Rightarrow s(u, t)).$$

Show that there is no relation on $M_{\{0\}}$ which, taken as ψa, makes $M_{\{0\}}$ a model of $\mathscr{A}$. Hence show that addition is not definable in $\mathscr{P}$.

4.10. Show that not every model of $\mathscr{P}$ is embeddable in a model of $\mathscr{S}(ZF$ set theory).

4.11. Taking $x \leqslant y$ as an abbreviation for $(\exists z)a(x, y, z)$, show that the axiom$ of a total order are theorems of $\mathscr{A}$.

Chapter VII

Ultraproducts

§1 Ultraproducts

In many branches of mathematics, where one is studying a system of some particular type, it is of interest to find out ways of forming new systems of the given type from known examples. One useful method that can often be applied is based on the cartesian product construction. In this section we investigate this construction in the case where the underlying system is a first-order theory $\mathscr{T} = (\mathscr{R}, A, C)$, and (M_i, v_i, ψ_i) for $i \in I$ is a family of models of $\mathscr{T}$. We therefore investigate the possibility of making $M = \prod_{i\in I} M_i$ into a model of $\mathscr{T}$, independently of the particular nature of $\mathscr{T}$.

An element of $\prod_{i\in I} M_i$ is a function $a: I \to \bigcup_{i\in I} M_i$ such that $a(i) \in M_i$. We shall when convenient denote $a(i)$ by a_i, and call it the i-component of a. There is now an obvious way to proceed. We define $v: C \to M$ by putting $v(c)_i = v_i(c)$, and we define ψr, for $r \in \mathscr{R}_n$, by putting $(a^{(1)}, \ldots, a^{(n)}) \in \psi r$ if $(a_i^{(1)}, \ldots, a_i^{(n)}) \in \psi_i r$ for all $i \in I$.

This construction gives a model M of $\mathscr{T}$ in some cases. For example, since a cartesian product of groups is a group, the method works for the case of elementary group theory. However, the method does not work in the case of elementary field theory, because a cartesian product of fields is a commutative ring with 1 having non-zero noninvertible elements. (This is easily seen, because all operations are defined componentwise, and hence $a \in M$ has an inverse if and only if each a_i is invertible. Take an a in which some but not all a_i are invertible.) Hence the above construction must be modified if it is to work for all theories $\mathscr{T}$. We shall have to define $\psi: \mathscr{R} \to \text{rel}(M)$ in such a way that for *every* $p(x_1, \ldots, x_n) \in P(V, \mathscr{R})$, the relation ψp given by p on M corresponds to the relations $\psi_i p$ given on the M_i in precisely the way that the ψr for $r \in \mathscr{R}$ correspond to the $\psi_i r$.

We simplify notation and work only with one variable formulae $p(x)$. (The n-variable case is covered by regarding x as an n-tuple $(x_1, \ldots, x_n)$.) We shall modify the definition of ψ by taking $a \in M$ to be in $\psi p(x)$ if $a_i \in \psi_i p(x)$ for all i in some "suitable" subset of I, where we have yet to decide which subsets of I are to be considered suitable. Since the definition is to apply to all $p \in P$, it applies to $\mathscr{I}(x, y)$. This means that if any subset other than I itself is allowed, $\psi\mathscr{I}$ will not be the identity relation on $\prod_{i\in I} M_i$, but merely an equivalence relation. Therefore we must reduce modulo $\psi\mathscr{I}$ in order to obtain a model of $\mathscr{T}$—the equivalence classes will be the elements of the model.

We now investigate the conditions a family of "suitable" subsets of I must satisfy. Denote such a family by $\mathscr{F}$. Let $p(x), q(x) \in P$, $a \in \prod_{i\in I} M_i$, and let $A = \{i \in I | a_i \text{ satisfies } p(x)\}$, $B = \{i \in I | a_i \text{ satisfies } q(x)\}$. Since a formula should hold for some i if it is to hold at all, we have

(i) $\varnothing \notin \mathscr{F}$.

If $A \in \mathscr{F}$, then a satisfies $p(x)$ and so must satisfy $p(x) \vee q(x)$, whatever $q(x)$ may be. Thus for any $B \subseteq I$, $A \cup B \in \mathscr{F}$ if $A \in \mathscr{F}$. Hence

(ii) Every subset of I which contains a set of $\mathscr{F}$ belongs to $\mathscr{F}$.

If $A \in \mathscr{F}$ and $B \in \mathscr{F}$, then a satisfies $p(x)$ and $q(x)$ and so must satisfy $p(x) \wedge q(x)$. Thus $A \cap B \in \mathscr{F}$ if $A, B \in \mathscr{F}$. Generalising to finite subfamilies of $\mathscr{F}$, we have

(iii) Every finite intersection of sets of $\mathscr{F}$ belongs to $\mathscr{F}$.

Finally, since a must satisfy exactly one of $p(x)$ or $\sim p(x)$, we have

(iv) For each $A \subseteq I$, exactly one of A and $I - A$ belongs to $\mathscr{F}$.

Definition 1.1. A set $\mathscr{F}$ of subsets of I satisfying the conditions (i), (ii) and (iii) above is called a *filter* on I. A filter which satisfies (iv) is called an *ultrafilter*.

The filters on I, being subsets of Pow(I), are partially ordered by inclusion. The ultrafilters are the maximal elements of the set of filters.

Examples

1.2. If $I \neq \varnothing$, $\{I\}$ is a filter on I.

1.3. If k is a fixed element of I, $F = \{J \subseteq I | k \in J\}$ is an ultrafilter on I. (Ultrafilters constructed in this way are called principal ultrafilters.)

1.4. If I is infinite, the complements of the finite subsets of I form a filter. (When $I = \mathbf{N}$, this filter is called the Fréchet filter.)

Exercise 1.5. $\mathscr{F}$ is an ultrafilter on I and $J \in \mathscr{F}$. Prove that $\mathscr{F}_J = \{A \cap J | A \in \mathscr{F}\}$ is an ultrafilter on J, and that for $A \subseteq I$, $A \in \mathscr{F}$ if and only if $A \cap J \in \mathscr{F}_J$. ($\mathscr{F}_J$ is called the restriction of $\mathscr{F}$ to J.)

Let $a, b \in \prod_{i \in I} M_i$ and let $\mathscr{F}$ be an ultrafilter on I. We write $a \equiv b \bmod \mathscr{F}$ if $\{i \in I | a_i = b_i\} \in \mathscr{F}$, and denote the congruence class containing a by $a\mathscr{F}$. The set of all congruence classes is denoted by $\prod_{i \in I} M_i/\mathscr{F}$. For each $r \in \mathscr{R}$, we define the relation ψr on $\prod_{i \in I} M_i/\mathscr{F}$ by $a\mathscr{F} \in \psi r$ if $\{i \in I | a_i \in \psi_i r\} \in \mathscr{F}$. (Here, a is an n-tuple if $r \in \mathscr{R}_n$.) This definition is clearly independent of the choice of representative of the congruence class. To complete the construction, we define $v(c)$ for $c \in C$ to be the congruence class of the function $I \to \bigcup_{i \in I} M_i$ whose i-component is $v_i(c)$.

Theorem 1.6. *$\prod_{i \in I} M_i/\mathscr{F}$ is a model of $\mathscr{T} = (\mathscr{R}, A, C)$. An element $a\mathscr{F}$ of $\prod_{i \in I} M_i/\mathscr{F}$ satisfies $p(x) \in P$ (where a, x may be n-tuples) if and only if $\{i \in I | a_i$ satisfies $p(x)\} \in \mathscr{F}$.*

Proof: $\prod_{i \in I} M_i/\mathscr{F}$ is clearly a model of $\mathscr{T}' = (\mathscr{R}, \varnothing, C)$. To show that it is a model of $\mathscr{T}$, we have to show that $v(p) = 1$ for all $p \in A$. Since for $p \in A$, $\{i \in I | p$ is true in $M_i\} = I \in \mathscr{F}$, this will be an immediate consequence of the second assertion of the theorem. We shall prove this latter assertion by induction over the length of p.

If $p = r(x)$, where $r \in \mathscr{R}$, then the result holds by the definition of ψr. If $p = q_1 \Rightarrow q_2$, then $v(p) = 0$ if and only if we have $v(q_1) = 1$ and $v(q_2) = 0$. By induction, this holds precisely when $J_1 = \{i \in I | v_i(q_1) = 1\}$ and $J_2 = \{i \in I | v_i(q_2) = 0\}$ are both in $\mathscr{F}$. Put $J_3 = J_1 \cap J_2$. If $J_3 \in \mathscr{F}$, then $J_1 \in \mathscr{F}$ and $J_2 \in \mathscr{F}$ by condition (ii), while $J_1 \in \mathscr{F}$ and $J_2 \in \mathscr{F}$ imply $J_3 \in \mathscr{F}$ by condition (iii). Thus $v(p) = 0$ if and only if $J_3 = \{i \in I | v_i(p) = 0\} \in \mathscr{F}$. By condition (iv), $v(p) = 1$ if and only if $I - J_3 = \{i \in I | v_i(p) = 1\} \in \mathscr{F}$.

If $p(x) = (\forall y)q(x, y)$, then $a\mathscr{F}$ satisfies $p(x)$ if and only if for every $b\mathscr{F} \in \prod_{i \in I} M_i/\mathscr{F}$, $(a\mathscr{F}, b\mathscr{F})$ satisfies $q(x, y)$. By induction, the latter holds if and only if for all $b\mathscr{F}$, $\{i \in I | (a_i, b_i)$ satisfies $q(x, y)\} \in \mathscr{F}$. Let $J = \{i \in I | a_i$ satisfies $p(x)\}$. Suppose $J \in \mathscr{F}$. Then for all $i \in J$ and all $b\mathscr{F}$, we have (a_i, b_i) satisfies $q(x, y)$ since a_i satisfies $(\forall y)q(x, y)$. Thus $a\mathscr{F}$ satisfies $p(x)$. Suppose $J \notin \mathscr{F}$. Then for each $i \in K = I - J$, there exists an element $b_i \in M_i$ such that (a_i, b_i) does not satisfy $q(x, y)$. Thus there exists $b \in \prod_{i \in I} M_i$ such that, for all $i \in K$, (a_i, b_i) does not satisfy $q(x, y)$. Since $K \in \mathscr{F}$, $(a\mathscr{F}, b\mathscr{F})$ does not satisfy $q(x, y)$ and $a\mathscr{F}$ does not satisfy $p(x)$. □

Definition 1.7. The model $\prod_{i \in I} M_i/\mathscr{F}$ of $\mathscr{T}$ is called the *ultraproduct* of the models M_i with respect to the ultrafilter $\mathscr{F}$.

Exercises

1.8. Let p_i be the ith prime and let F_i be a field of characteristic p_i. Let $\mathscr{F}$ be an ultrafilter on the set I of positive integers, such that no member of $\mathscr{F}$ is a singleton. Prove that $\prod_{i \in I} F_i/\mathscr{F}$ is a field of characteristic zero.

1.9. $\mathscr{F}$ is an ultrafilter on I, M_i $(i \in I)$ are models of the theory $\mathscr{T}$, and $J \in \mathscr{F}$. Prove

$$\prod_{i \in I} M_i/\mathscr{F} \simeq \prod_{j \in J} M_j/\mathscr{F}_J,$$

where $\mathscr{F}_J$ is the restriction of $\mathscr{F}$ to J.

§2 Non-Principal Ultrafilters

Principal ultrafilters on I, as constructed in Exercise 1.3, are of no use for the construction of new models, because an ultraproduct with respect to a principal ultrafilter is always isomorphic to one of the factors.

Exercises

2.1. If $k \in I$ and $\mathscr{F} = \{J \subseteq I | k \in J\}$, prove that

$$\prod_{i \in I} M_i/\mathscr{F} \simeq M_k.$$

2.2. $\mathscr{F}$ is an ultrafilter on I and $A \in \mathscr{F}$ is a finite subset of I. Prove that $\mathscr{F}$ is principal.

We now investigate conditions on a set S of subsets of I for the existence of an ultrafilter $\mathscr{F} \supseteq S$. By an appropriate choice of S, we shall be able to ensure that every such ultrafilter is non-principal.

Definition 2.3. The set S of subsets of I is said to have the *finite intersection property* if every finite subset of S has non-empty intersection.

Lemma 2.4. *Let S be a set of I. There exists a filter on I containing S if and only if S has the finite intersection property.*

Proof: The necessity of the condition is immediate, so we prove its sufficiency. Suppose S has the finite intersection property, and put

$$T = \{U \subseteq I \mid U = J_1 \cap \cdots \cap J_n \text{ for some } n \text{ and some } J_1, \ldots, J_n \in S\}.$$

Let

$$\mathscr{F} = \{F \subseteq I \mid F \supseteq U \text{ for some } U \in T\}.$$

We prove that $\mathscr{F}$, which clearly contains S, is a filter. By the finite intersection property of S, $\varnothing \notin T$ and so $\varnothing \notin \mathscr{F}$. Also, condition (ii) for a filter is clearly satisfied by $\mathscr{F}$. Finally, if $F_1, \ldots, F_n \in \mathscr{F}$, then for $i = 1, \ldots, n$, $F_i \supseteq \bigcap_{j=1}^{m_i} J_{ij}$ for some m_i and $J_{i1}, \ldots, J_{im_i} \in S$. Hence.

$$\bigcap_{i=1}^{n} F_i \supseteq \bigcap_{i=1}^{n} \bigcap_{j=1}^{m_i} J_{ij},$$

and so belongs to $\mathscr{F}$. Thus condition (iii) is satisfied and $\mathscr{F}$ is a filter. □

Lemma 2.5. *Let $\mathscr{F}$ be a filter on I. Then there exists an ultrafilter $\mathscr{F}^* \supseteq \mathscr{F}$ on I.*

Proof: The set of filters containing $\mathscr{F}$ is an inductive set. By Zorn's Lemma, it has a maximal member $\mathscr{F}^*$. □

Exercises

2.6. Let $\alpha = |I|$ and suppose $\alpha \geqslant \beta \geqslant \aleph_0$. Put $S = \{J \subseteq I \mid |I - J| < \beta\}$. Prove that S is a filter and that if $\mathscr{F}$ is an ultrafilter containing S, then no member of $\mathscr{F}$ has cardinal less than β.

2.7. An ultrafilter $\mathscr{F}$ on I is called uniform if $|J| = |I|$ for all $J \in \mathscr{F}$. If $\mathscr{F}$ is a non-principal ultrafilter, show that there exists $J \in \mathscr{F}$ such that $\mathscr{F}_J$ is uniform.

2.8. Let I be a countable set, and $\mathscr{F}$ an untrafilter on I. If $\sigma: I \to I$ is a permutation, show that $\sigma\mathscr{F}$ is also an ultrafilter on I. The collection $\{\sigma\mathscr{F} \mid \sigma \text{ a permutation of } I\}$ may be called the orbit of $\mathscr{F}$. Show that if $\mathscr{F}$ is non-principal, its orbit contains exactly $2^{\aleph_0}$ distinct ultrafilters.

2.9. A family $\mathscr{A}$ of infinite subsets of an infinite set X is called almost disjoint (AD) if distinct members of $\mathscr{A}$ have finite intersection. $\mathscr{A}$ is called maximal almost disjoint (MAD) if it is maximal among the AD families. Prove or disprove each of the following:

(a) Given any MAD family $\mathscr{A}$, there is a non-principal ultrafilter $\mathscr{F}$ such that $\mathscr{A}$ and $\mathscr{F}$ are disjoint.

(b) Given any non-principal ultrafilter $\mathscr{F}$ there is a MAD family $\mathscr{A}$ such that $\mathscr{A}$ and $\mathscr{F}$ are disjoint.

§3 The Existence of an Algebraic Closure

We can now apply the theory of ultraproducts to prove a theorem of considerable importance in algebra.

Theorem 3.1. *Let F be a field. Then there exists an algebraic closure of F.*

Proof: Let $\mathscr{T}$ be elementary field theory augmented by the addition of the elements of F to the set of constants, and of all the relations $a_1 + a_2 = a_3$, $b_1 b_2 = b_3$ holding in F to the set of axioms. The models of $\mathscr{T}$ are the extension fields of F. Put $R = F[x]$, the ring of polynomials over F. For each $r \in R$, let F_r be a splitting field of r. Put

$$J_r = \{s \in R \mid r \text{ splits over } F_s\}.$$

Since $r_1 r_2 \cdots r_n \in J_{r_1} \cap J_{r_2} \cap \cdots \cap J_{r_n}$, the set $\mathscr{J} = \{J_r \mid r \in R\}$ has the finite intersection property. By Lemmas 2.4 and 2.5, there exists an ultrafilter $\mathscr{F}$ on R containing $\mathscr{J}$. Put $F^* = \prod_{r \in R} F_r / \mathscr{F}$. Then F^* is a model of $\mathscr{T}$ and so is an extension field of $\mathscr{F}$.

Let $r = x^n + r_1 x^{n-1} + \cdots + r_n$ be a monic polynomial over F. We prove that r splits over F^*. We put

$$p = (\exists a_1) \cdots (\exists a_n)((a_1 + \cdots + a_n = -r_1) \wedge (a_1 a_2 + a_1 a_3 + \cdots + a_{n-1} a_n = r_2) \wedge \cdots \wedge (a_1 a_2 \cdots a_n = (-1)^n r_n)).$$

Then p is true for precisely those models of $\mathscr{T}$ over which r splits. But $\{s \in R \mid p \text{ is true in } F_s\} = J_r \in \mathscr{F}$. By Théorem 1.6, p is true in F^* and so r splits over F^*.

The proof of Theorem 3.1 is completed by the following purely algebraic lemma.

Lemma 3.2. *Let F^* be an extension of the field F such that every monic polynomial over F splits over F^*. Let $\bar{F}$ be the set of all elements of F^* which are algebraic over F. Then $\bar{F}$ is an algebraic closure of F.*

Proof: Let $f(x)$ be a monic polynomial over $\bar{F}$. Then $f(x) = (x - a_1) \cdots (x - a_n)$ for some $a_1, \ldots, a_n$ in the splitting field of $f(x)$ considered as a polynomial over F^*. But the a_i, being algebraic over $\bar{F}$, are algebraic over F. Let $m_i(x)$ be the minimum polynomial of a_i over F. Since $m_i(x)$ splits over F^*, its roots lie in F^* and, being algebraic over F, are therefore in $\bar{F}$. Thus $a_1, \ldots, a_n \in \bar{F}$ and $f(x)$ splits over $\bar{F}$. Hence $\bar{F}$ is algebraically closed. □

Exercises

3.3. If F is not algebraically closed, prove that the ultrafilter used in the proof of Theorem 3.1 is not principal.

3.4. In the notation of the proof of Theorem 3.1, show that if F is finite, then $\{s \in R | F_s = F_r\} \notin \mathscr{F}$.

3.5. If F is not algebraically closed, prove that F^* (constructed as above) is not algebraic over F. (If F is infinite, show that elements $a_r \in F_r$ can be chosen such that a_r and a_s have the same minimum polynomial only for $r = s$. If F is finite, show that the elements $a_r \in F_r$ can be chosen such that a_r and a_s have the same minimum polynomial only for those r, s for which $F_r = F_s$.)

3.6. F is a field. For all $i \in I$, take $F_i = F$ and form the ultraproduct $K = \prod_{i \in I} F_i/\mathscr{F}$ with respect to the ultrafilter $\mathscr{F}$. Prove that K is a pure transcendental extension of F.

§4 Non-trivial Ultrapowers

An ultraproduct $\prod_{i \in I} M_i/\mathscr{F}$ in which $M_i = M$ for all $i \in I$ is called an *ultrapower* of M and denoted by $M^I/\mathscr{F}$. There is a natural embedding $\theta: M \to M^I/\mathscr{F}$ of M in $M^I/\mathscr{F}$ given by $\theta(m) = f_m\mathscr{F}$, where $f_m: I \to M$ is the constant function $f_m(i) = m$ for all $i \in I$. By identifying m with $\theta(m)$, we may regard M as a subset of $M^I/\mathscr{F}$. (Alternatively, we may replace the theory $\mathscr{T}$ by the theory $\mathscr{T}'$ formed from $\mathscr{T}$ by replacing C by $C \cup M$. By Theorem 1.6, $M^I/\mathscr{F}$ is a model of $\mathscr{T}'$. Since each element $m \in M$ is a constant of $\mathscr{T}'$, this also gives a map $v': M \to M^I/\mathscr{F}$.)

Exercise 4.1. Prove that the maps $\theta, v': M \to M^I/\mathscr{F}$ coincide.

We shall always make this identification of M with $\theta(M)$, and we omit specific mention of the map θ. The ultrapower $M^I/\mathscr{F}$ is regarded as trivial if $M^I/\mathscr{F} = M$, so we shall look for conditions which ensure non-triviality.

Exercise 4.2. If M is finite, prove that $M^I/\mathscr{F} = M$.

Definition 4.3. Let α be a cardinal. The ultrafilter $\mathscr{F}$ on I is called *α-complete* if, for every subset $\mathscr{G} \subseteq \mathscr{F}$ of cardinal α, we have $\cap\mathscr{G} \in \mathscr{F}$. Otherwise, $\mathscr{F}$ is called *α-incomplete.* (It is usual in this context to denote $|\mathbb{N}|$ by ω.)

Lemma 4.4. *Let α be an infinite cardinal and let $\mathscr{F}$ be an α-incomplete ultrafilter on I. Then there exists a partition of I into α disjoint subsets, none of which is in $\mathscr{F}$.*

Proof: The cardinal α is an ordinal, $\alpha = \{\beta | \beta \text{ ordinal}, \beta < \alpha\}$. Since $\mathscr{F}$ is α-incomplete, there exists $\mathscr{G} \subseteq \mathscr{F}$ such that $|\mathscr{G}| = \alpha$ and $\cap\mathscr{G} \notin \mathscr{F}$. We index the members of $\mathscr{G}$ with the ordinals less than α, so that $\mathscr{G} = \{G_\beta | \beta < \alpha\}$. For each ordinal $\beta \leqslant \alpha$, put $X_\beta = \cap\{G_\gamma | \gamma < \beta\}$ (interpreting this for $\beta = 0$ to mean $X_0 = I$), and put $Y_\beta = X_\beta - X_{\beta+1}$ for $\beta < \alpha$. For $\beta = \alpha$, put

$Y_\alpha = X_\alpha$. Then $\{Y_\beta | \beta \leqslant \alpha\}$ is a partition of I into α disjoint subsets. Since $Y_\alpha = \cap\mathscr{G}$, we have $Y_\alpha \notin \mathscr{F}$. Suppose $Y_\beta \in \mathscr{F}$ for some $\beta < \alpha$. Then $X_\beta - X_{\beta+1} \in \mathscr{F}$. Since also $G_\beta \in \mathscr{F}$, we have $(X_\beta - X_{\beta+1}) \cap G_\beta \in \mathscr{F}$. But $(X_\beta - X_{\beta+1}) \cap G_\beta = \varnothing \notin \mathscr{F}$. □

Lemma 4.5. *Let $\mathscr{F}$ be an α-complete ultrafilter on I. Then for every partition of I into a set $\mathscr{G}$ of α disjoint subsets, some member of $\mathscr{G}$ is in $\mathscr{F}$.*

Exercise 4.6. Prove Lemma 4.5.

Theorem 4.7. *Let $\mathscr{F}$ be an ultrafilter on I and let $\alpha = |M|$. Then $M = M^I/\mathscr{F}$ if and only if $\mathscr{F}$ is α-complete.*

Proof: Suppose $\mathscr{F}$ is α-complete. An element of $M^I/\mathscr{F}$ is $f\mathscr{F}$ for some $f: I \to M$. For each $m \in M$, put $J_m = \{i \in I | f(i) = m\}$. Then $\{J_m | m \in M\}$ is a partition of I into α disjoint subsets. By Lemma 4.6, $J_m \in \mathscr{F}$ for some $m \in M$. This implies $f\mathscr{F} = m$.

Suppose now that $\mathscr{F}$ is α-incomplete. Then α must be infinite, and so, by Lemma 4.4, there is a partition of I into α disjoint subsets, none of which is in $\mathscr{F}$. We may index these subsets with the elements of M. Let $\{J_m | m \in M\}$ be such a partition of I, and let $f(i)$ be the unique $m \in M$ such that $i \in J_m$. This defines a function $f: I \to M$ such that $f\mathscr{F} \notin M$. □

Exercise 4.8. $\mathscr{F}$ is a non-principal ultrafilter on I, and β is the smallest cardinal for which $\mathscr{F}$ is β-incomplete. $|M| = \alpha \geqslant \beta$. Prove that $|M^I/\mathscr{F}| \geqslant \alpha^\beta$.

It can be proved (cf [1], p. 112, Theorem 1.11) that if $|I|$ is less than the first strongly inaccessible cardinal, then every non-principal ultrafilter on I is ω-incomplete. This means that if $\mathscr{F}$ is non-principal, then $M^I/\mathscr{F} \neq M$ except when I is very large or when M is finite.

Exercise 4.9. Let α be an infinite cardinal and let A be a set of cardinal α. Put $M = A \cup \text{Pow}(A)$, and $\mathscr{R} = \{\mathscr{I}, \in, e, s\}$ where e, s are unary and $\mathscr{I}, \in$ are binary. Interpreting $e(x)$ as x is an element of A, $s(x)$ as x is a subset of A, and $\in(x, y)$ as x is a member of y (for x in A, y in $\text{Pow}(A)$), form the theory $\mathscr{T} = (\mathscr{R}, A, C)$ with $\mathscr{R}$ as above, $C = M$ and $A = \{p \in \mathscr{L}(\mathscr{T}) | p \text{ true in } M\}$. For any model N of $\mathscr{T}$, put $B = \{n \in N | e(n) \text{ is true}\}$ and $D = \{n \in N | s(n) \text{ is true}\}$. Show that each $d \in D$ is determined by the set $\{b \in B | \in(b, d) \text{ is true}\}$ and hence identify D with a subset of $\text{Pow}(B)$. In the special case where $\mathscr{F}$ is an α-complete ultrafilter on a set I of cardinal 2^α and $N = M^I/\mathscr{F}$, prove that $B = A$ and $N = M$. Hence prove that an α-complete ultrafilter on a set of cardinal 2^α is principal.

§5 Ultrapowers of Number Systems

We have seen that the theory of N (i.e., the theory $\mathscr{T} = (\mathscr{R}, A, C)$ where $\mathscr{R} = \{\mathscr{I}, +, \times, <\}$, $C = \text{N}$ and $A = \{p \in \mathscr{L}(\mathscr{T}) | p \text{ true in N}\}$ cannot be categorical, and the same is true for the theories of the other standard systems Z, Q, R and C. We use ultrapowers to produce models of these theories which are not isomorphic to their standard models. We take the set N as index set I. Let $\mathscr{F}$ be a non-principal ultrafilter on I. Since $\mathscr{F}$ contains no finite sets, every subset of I with finite complement is in $\mathscr{F}$ (i.e., $\mathscr{F}$ is an extension of the Fréchet filter on I). Trivially, $\mathscr{F}$ is ω-incomplete. By Theorem 4.7, if M is any of N, Z, Q, R or C, then $M^I/\mathscr{F} \neq M$.

An element of N^I is just a sequence of natural numbers. When we form $\mathrm{N}^I/\mathscr{F}$, we are, among other things, identifying sequences which are the same from some point onwards. Consider the element $u\mathscr{F}$ of $\mathrm{N}^I/\mathscr{F}$ given by the function $u: \mathrm{N} \to \mathrm{N}$ defined by $u(i) = i$. This element $u\mathscr{F}$ is infinite, in the sense that $u\mathscr{F} > n$ for all $n \in \mathrm{N}$. To see this, let $k_n: \mathrm{N} \to \mathrm{N}$ be the constant function $k_n(i) = n$ for all $i \in \mathrm{N}$. Then $n = k_n\mathscr{F}$ and $\{i \in \mathrm{N} | u(i) > k_n(i)\} = \{i \in \mathrm{N} | i > n\} \in \mathscr{F}$. Hence by Theorem 1.6, $u\mathscr{F} > k_n\mathscr{F}$. Similarly, we can show that $\mathrm{Q}^I/\mathscr{F}$ has infinitesimal elements. For if $v: \mathrm{N} \to \mathrm{Q}$ is defined by $v(i) = 1/i$ for $i > 0$, $v(0) = 1$, then $k_0\mathscr{F} < v\mathscr{F} < k_r\mathscr{F}$ for all $r \in \mathrm{Q}$ such that $r > 0$. We clearly have natural inclusions $\mathrm{N}^I/\mathscr{F} \subseteq \mathrm{Z}^I/\mathscr{F} \subseteq \mathrm{Q}^I/\mathscr{F} \subseteq \mathrm{R}^I/\mathscr{F} \subseteq \mathrm{C}^I/\mathscr{F}$.

Exercise 5.1. Let N be a model of the theory of N which properly contains N. Show that N has infinite elements. Show also that if Q is a model of the theory of Q which properly contains Q, then Q has non-zero infinitesimal elements.

Let $\mathscr{T}$ be the theory of N. Form the theory $\mathscr{T}'$ by adding a new constant u and the new axioms $u > n$ for all $n \in \mathrm{N}$. This theory $\mathscr{T}'$ is consistent, indeed $\mathrm{N}^I/\mathscr{F}$, with u interpreted as $u\mathscr{F}$, is a model of $\mathscr{T}'$. As $\mathscr{T}'$ is a countable theory, the Löwenheim-Skolem Theorem (Theorem 3.18 of Chapter V) shows that it has a countable model. The model $\mathrm{N}^I/\mathscr{F}$ is uncountable, and it is natural to try to modify the ultrapower construction so as to obtain a countable model. We shall take a subset S of the set N^I of all functions from I into N and reduce this set S modulo an ultrafilter $\mathscr{F}$.

Let $\mathscr{T} = (\mathscr{R}, A, C)$ be any theory, $\{M_i | i \in I\}$ a family of models of $\mathscr{T}$, and $\mathscr{F}$ an ultrafilter on I. If S is any non-empty subset of $\prod_{i\in I} M_i$, which includes all the functions $k_c: I \to \prod_{i\in I} M_i$ defined by $k_c(i) = v_i(c)$ for $c \in C$, $i \in I$, then $S/\mathscr{F}$ is a model of $(\mathscr{R}, \varnothing, C)$.

Definition 5.2. $S/\mathscr{F}$ is called a *subultraproduct* of $\{M_i | i \in I\}$ with respect to the ultrafilter $\mathscr{F}$.

A subultraproduct $S/\mathscr{F}$ of models M_i of $\mathscr{T} = (\mathscr{R}, A, C)$ is a model of $(\mathscr{R}, \varnothing, C)$, but unless further conditions are imposed on S or on A, it need not be a model of $\mathscr{T}$. If we examine the proof of Theorem 1.6, we see that it applies unaltered, except for the section which shows that $a\mathscr{F}$ satisfies $p(x) = (\forall y)q(x, y)$ only if the set $J = \{i \in I | a_i \text{ satisfies } p(x)\}$ is in $\mathscr{F}$. For each $i \in I - J$, there exists an element $b_i \in M_i$ such that (a_i, b_i) does not satisfy $q(x, y)$, but we can no longer conclude from this that there exists a function $b \in S$ such that for all $i \in I - J$, (a_i, b_i) does not satisfy $q(x, y)$. If S is chosen so that these functions always exist, then the assertions of Theorem 1.6 will continue to hold for $S/\mathscr{F}$. The next theorem shows how to achieve this. The reader is asked to recall Definition 2.11 of Chapter V.

Theorem 5.3. *Let $\mathscr{T}$ be the theory of* N. *Let S be the set of all functions $s: \mathrm{N} \to \mathrm{N}$ which are definable in $\mathscr{T}$. Let $\mathscr{F}$ be an ultrafilter on* N. *Then the subultraproduct $S/\mathscr{F}$ is a countable model of $\mathscr{T}$. If $\mathscr{F}$ is not principal, then $S/\mathscr{F} \neq \mathrm{N}$.*

Proof: To show that $S/\mathscr{F}$ is a model of $\mathscr{T}$, it remains for us to show that if $a \in S$ and $J = \{i \in \mathrm{N} | a_i \text{ satisfies } (\forall y)q(x, y)\} \notin \mathscr{F}$, then there exists $b \in S$ such that, for all $i \in \mathrm{N} - J$, we have that (a_i, b_i) does not satisfy $q(x, y)$. We put

$$\begin{aligned} b_i &= 0 \text{ if } a_i \text{ satisfies } (\forall y)q(x, y), \\ b_i &= \text{the least } n \text{ for which } (a_i, n) \text{ does not satisfy } q(x, y) \\ &\quad \text{if } a_i \text{ does not satisfy } (\forall y)q(x, y). \end{aligned}$$

Since a is a definable function, so is b, and the assertion follows. Since P is countable, so is S and hence so is $S/\mathscr{F}$. The function $u: \mathrm{N} \to \mathrm{N}$ given by $u(i) = i$ is clearly definable, as are the constant functions k_n. If $\mathscr{F}$ is non-principal, then $\{i \in \mathrm{N} | u(i) = k_n(i)\} = \{n\} \notin \mathscr{F}$ and so $S/\mathscr{F} \neq \mathrm{N}$. □

§6 Direct Limits

There is a connection between the idea of a subultraproduct, introduced in §5, and the idea of a direct limit, which arises in a number of algebraic contexts, and which we now explain. A directed set is a partially ordered set $(I, \leqslant)$ such that for any $i, j \in I$, there is a $k \in I$ such that $i \leqslant k$ and $j \leqslant k$. A direct family[1] in a category $\mathscr{C}$ is a set $\{A_i | i \in I\}$ of objects A_i of $\mathscr{C}$ indexed by a directed set I, together with a morphism $f^i_j: A_i \to A_j$ for each pair $i \leqslant j$ in I, such that

(i) $f^i_i = 1_{A_i}$ for all $i \in I$,
(ii) $f^j_k f^i_j = f^i_k$ for all $i \leqslant j \leqslant k$ in I.

Definition 6.1. A *direct limit* in $\mathscr{C}$, or more precisely a *limit of the direct family* $\{A_i, f^i_j | i, j \in I\}$ in $\mathscr{C}$, is an object L of $\mathscr{C}$ together with morphisms $\varphi^i: A_i \to L$ such that

(i) $\varphi^j f^i_j = \varphi^i$ for all $i \leqslant j$ in I, and
(ii) for any object M of $\mathscr{C}$ and family of morphisms $\psi^i: A_i \to M$ satisfying $\psi^j f^i_j = \psi^i$ for all $i \leqslant j$ in I, there exists a unique morphism $\theta: L \to M$ such that $\theta\varphi^i = \psi^i$ for all $i \in I$.

[1] More neatly but less intuitively defined as follows: regard the directed set as a category with objects the elements of I and morphisms from i to j the pairs (i, j) with $i \leqslant j$. A direct family in $\mathscr{C}$ is then a functor from this category into $\mathscr{C}$.

Let $\mathscr{T} = (\mathscr{R}, A, C)$ be a theory. We associate with $\mathscr{T}$ the category $\mathrm{Mod}(\mathscr{T})$, whose objects are the models of $\mathscr{T}$ and whose morphisms are the maps $f: M \to M'$ between models (M, v, ψ) and (M', v', ψ'), such that $fv(c) = v'(c)$ for all $c \in C$, and such that $(f(m_1), \dots, f(m_n)) \in \psi' r$ for all $r \in \mathscr{R}_n$ and $(m_1, \dots, m_n) \in \psi r$. As an aid to the study of $\mathrm{Mod}(\mathscr{T})$, we use the theory $\mathscr{T}_\varnothing = (\mathscr{R}, \varnothing, C)$ and its associated category $\mathrm{Mod}(\mathscr{T}_\varnothing)$. Every object of $\mathrm{Mod}(\mathscr{T})$ is an object of $\mathrm{Mod}(\mathscr{T}_\varnothing)$, and the morphisms in $\mathrm{Mod}(\mathscr{T})$ between any two of its objects are precisely the morphisms between them in $\mathrm{Mod}(\mathscr{T}_\varnothing)$.

Exercise 6.2. R is a ring and $\mathscr{T}$ is the elementary theory of R-modules. Show that the category of R-modules is precisely the category $\mathrm{Mod}(\mathscr{T})$.

Lemma 6.3. *For any theory $\mathscr{T}$, direct limits exist in* $\mathrm{Mod}(\mathscr{T}_\varnothing)$.

Proof: Let $\{M_i, f_j^i | i, j \in I, i \leqslant j\}$ be a direct family in $\mathrm{Mod}(\mathscr{T}_\varnothing)$. We construct a limit as a subultraproduct. Put $F_i = \{j \in I | j \geqslant i\}$, and let $\mathscr{F}$ be any ultrafilter containing all the F_i. (Actually, we do not need $\mathscr{F}$ to be an ultrafilter—any filter containing the F_i will do.) Put

$$S = \{s: I \to \bigcup_{i \in I} M_i | s_i \in M_i \text{ for all } i \in I; \text{ for some } F \in \mathscr{F}, f_j^i s_i = s_j \text{ for all } i, j \in F \text{ with } i \leqslant j\}.$$

Then put $L = S/\mathscr{F}$ and define $\varphi^i: M_i \to L$ by $\varphi^i(m) = s\mathscr{F}$, where s is given by $s_j = f_j^i m$ for all $j \in F_i$, and s_j is chosen arbitrarily for $j \notin F_i$. The element $s\mathscr{F}$ of L is clearly independent of the choice of s_j for $j \notin F_i$. L, being a subultraproduct of models M_i of $\mathscr{T}$, is a model of $\mathscr{T}_\varnothing$. The φ^i are clearly morphisms of $\mathrm{Mod}(\mathscr{T}_\varnothing)$ satisfying (i) of Definition 6.1.

Now let N be a model of $\mathscr{T}_\varnothing$, and let $\psi^i: M_i \to N$ be morphisms in $\mathrm{Mod}(\mathscr{T}_\varnothing)$ such that $\psi^j f_j^i = \psi^i$ for all $i \leqslant j$ in I. Let $s \in S$ satisfy $f_j^i s_i = s_j$ for all $i, j \in F$ such that $i \leqslant j$, where F is a member of $\mathscr{F}$. If $i \in F$ and if $m = s_i \in M_i$, then $s\mathscr{F} = \varphi^i(m)$. The condition on the map $\theta: L \to N$ to be constructed requires that $\theta(s\mathscr{F}) = \psi^i(m)$. Hence θ, if it exists, is unique. We define θ by putting $\theta(s\mathscr{F}) = \psi^i(s_i)$ for some $i \in F$, and we must show that this definition is independent of the choice of i.

Suppose that $j \in F$. Then there exists $k \in I$ such that $i \leqslant k$ and $j \leqslant k$. Since $F \cap F_k \in \mathscr{F}$, $F \cap F_k \neq \varnothing$, and so there exists an $r \in F \cap F_k$. We have $s_r = f_r^i s_i = f_r^j s_j$, and so

$$\psi^i s_i = \psi^r f_r^i s_i = \psi^r s_r = \psi^j s_j.$$

Thus θ is well-defined. Clearly, θ is a morphism satisfying the requirements of condition (ii) of the definition. □

We are interested in direct limits in $\mathrm{Mod}(\mathscr{T})$. The next lemma reduces this problem to an investigation of the subultraproduct constructed above.

Lemma 6.4. *The direct family $\{M_i, f_j^i | i, j \in I, i \leqslant j\}$ has a limit in* $\mathrm{Mod}(\mathscr{T})$ *if and only if the subultraproduct $S/\mathscr{F}$ constructed above is a model of $\mathscr{T}$.*

Proof: If $S/\mathscr{F}$ is a model of $\mathscr{T}$, then it is clearly a limit in Mod($\mathscr{T}$) of the given family. If L is a limit in Mod($\mathscr{T}$) of the family, then L is also a limit in Mod($\mathscr{T}_{\varnothing}$), and so is isomorphic to $S/\mathscr{F}$. Thus $S/\mathscr{F}$ is a model of $\mathscr{T}$. □

We now investigate conditions on $\mathscr{T}$ for $S/\mathscr{F}$ to be a model of $\mathscr{T}$.

Definition 6.5. An element $p \in P(V, \mathscr{R})$ is called *universal* if it has the form $(\forall y_1)\cdots(\forall y_s)q(x_1, \ldots, x_r, y_1, \ldots, y_s)$, where $q(x_1, \ldots, x_r, y_1, \ldots, y_s)$ contains no quantifiers.

The argument used in proving Theorem 1.6 shows that the element $s\mathscr{F}$ will satisfy $p(x) = (\forall y)q(x, y)$, where $q(x, y)$ contains no quantifiers, if $\{i \in I | s_i \text{ satisfies } p(x)\} \in \mathscr{F}$. (This includes the case where x, y are n-tuples.) Thus every axiom of $\mathscr{T}$ which is universal is satisfied in $S/\mathscr{F}$.

Definition 6.6. A theory $\mathscr{T}$ is called *algebraic* if every axiom of $\mathscr{T}$ is either universal or has the form

$$(\forall x_1)\cdots(\forall x_r)(\exists y_1)\cdots(\exists y_s)p(x_1, \ldots, x_r, y_1, \ldots, y_s, c_1, \ldots, c_t),$$

where $c_i \in C$ and p is constructed from primitive elements of $P(V, \mathscr{R})$ by using $\vee$, $\wedge$ only. The category Mod($\mathscr{T}$) is called *algebraic* if $\mathscr{T}$ is algebraic.

The reason for the name is that if the relation symbol $r \in \mathscr{R}_{n+1}$ is to correspond to an n-ary operation of an algebra, then we require the axioms $(\forall x_1)\cdots(\forall x_n)(\exists y)r(x_1,\ldots,x_n, y)$ and $(\forall x_1)\cdots(\forall x_n)(\forall y)(\forall z)(r(x_1,\ldots,x_n, y) \wedge r(x_1, \ldots, x_n, z) \Rightarrow y = z)$. Note that the second of these axioms is universal, and the first is admissible for an algebraic theory.

Theorem 6.7. *Let $\mathscr{T}$ be an algebraic theory. Then direct limits exist in* Mod($\mathscr{T}$).

Proof: Let $q = (\forall x)(\exists y)p(x, y, c)$ be an axiom of $\mathscr{T}$, where $p(x, y, c)$ is constructed from primitive elements of $P(V, \mathscr{R})$ using only $\vee$, $\wedge$. (x, y, c may denote n-tuples.) Let $a\mathscr{F} \in S/\mathscr{F}$ and let $\sigma_i = v_i(c)$. We have, for some $F \in \mathscr{F}$, $f^i_j a_i = a_j$ for all $i, j \in F$ with $i \leqslant j$. Take an $i \in F$, and put $F' = F \cap F_i$. Since q is satisfied in M_i, there exists $m_i \in M_i$ such that (a_i, m_i, σ_i) satisfies $p(x, y, c)$. For $j \in F'$, put $m_j = f^i_j m_i$. Since $a_j = f^i_j a_i$ and $\sigma_j = f^i_j \sigma_i$, it follows from the nature of p and the fact that f^i_j is a morphism that (a_j, m_j, σ_j) satisfies $p(x, y, c)$. Choose m_j arbitrarily for $j \notin F'$. Then $m \in S$ and $(a\mathscr{F}, m\mathscr{F}, \sigma\mathscr{F})$ satisfies $p(x, y, c)$. Hence q is satisfied in $S/\mathscr{F}$. □

Corollary 6.8. *Direct limits exist in any variety of universal algebras.*

Exercises

6.9. Show that direct limits exist in the category whose objects are fields and whose morphisms are ring homomorphisms, but that not even finite direct sums exist in this category. Show that the algebraic closure of a field is obtainable as a direct limit.

6.10. Show that the conditions we have imposed on the existential axioms cannot be weakened either by (a) allowing the negation of a primitive relation, or (b) allowing an existential quantifier to precede a universal quantifier. (Take the direct family, indexed by **N**, of the sets $\{0, 1, \ldots, n\}$ and inclusion maps, (a) with property $p(x)$ true in $\{0, 1, \ldots, n\}$ for $0, 1, \ldots, n - 1$, (b) with relation $\leqslant$ and axiom $(\exists x)(\forall y)(y \leqslant x)$.)

6.11. Show that compact topological spaces and continuous maps do not form an algebraic category.

Chapter VIII

Non-Standard Models

§1 Elementary Standard Systems

Much of mathematics is concerned with the study of "standard" mathematical systems such as the natural numbers, the rationals, the real numbers and the complex numbers, each of which is regarded as a unique system. When we attempt to study one of these systems by axiomatising it within the first-order predicate calculus, we find that our axiomatisation cannot be categorical, and that there exist models of our axiomatic theory not isomorphic to the system we wish to study. Such models have been constructed as ultrapowers in Chapter VII. In this chapter, we investigate ways of exploiting such models in the study of a standard system. We begin by considering elementary systems, i.e., systems in which relations between elements, but not properties of subsets, can be studied.

Definition 1.1. An *elementary standard system* S is a set S together with a subset $\mathscr{R}$ of the set of relations on S such that $\mathscr{I} \in \mathscr{R}$.

$\mathscr{R}$ is the set of relations on S considered to be of interest. It is usual to denote the underlying set S of S by the same symbol S, and we shall do so.

Example 1.2. The elementary real number system R consists of the set R of real numbers together with the set $\mathscr{R} = \{\mathscr{I}, +, \times, <\}$ of relations on R. Here, $+$ is the ternary relation $(a, b, c) \in +$ if and only if $a + b = c$, and $\times$ is defined similarly.

Let $\mathrm{S} = (\mathrm{S}, \mathscr{R})$ be an elementary standard system. We take a set $V \supset \mathrm{S}$, such that $V - \mathrm{S}$ is countably infinite, and form the first-order algebra $P(\mathrm{S}) = P(V, \mathscr{R})$. In this algebra, we think of elements of S and $\mathscr{R}$ as names for themselves[1].

Definition 1.3. The *language* of S is the subset $\mathscr{L}(\mathrm{S}) = \{p \in P(V, \mathscr{R}) | \operatorname{var}(p) \subseteq \mathrm{S}\}$ of $P(V, \mathscr{R})$.

Interpreting each element of S and $\mathscr{R}$ as itself assigns a truth value $v(p)$ to each $p \in \mathscr{L}(\mathrm{S})$.

Definition 1.4. The *theory* of S is the theory $\mathscr{T}(\mathrm{S}) = (\mathscr{R}, A, \mathrm{S})$ where $A = \{p \in \mathscr{L}(\mathrm{S}) | v(p) = 1\}$.

[1] If we wish to distinguish between the objects and their names, we take for each element $a \in \mathrm{S}$ and $\rho \in \mathscr{R}$ elements a', ρ', and use these in the construction of P.

$\mathscr{T}(\mathrm{S})$ is a complete theory with S as a model. The theorems of this theory are its axioms, and consist of all elements of $\mathscr{L}(\mathrm{S})$ which are true in S or in any other model of $\mathscr{T}(\mathrm{S})$. If the axiom set A were fully known, then $\mathscr{T}(\mathrm{S})$ could give us no new information about S. However, our knowledge of S is usually incomplete, and any method which extends our knowledge of A in fact extends our knowledge of S. If we can choose a model *S of $\mathscr{T}(\mathrm{S})$ such that the truth or falsity of certain statements $p \in \mathscr{L}(\mathrm{S})$ is more easily determined (by argument in the meta-language) for *S than for S, then we have a method of utilising $\mathscr{T}(\mathrm{S})$ to discover properties of S. Our aim is to construct some useful models *S.

Exercise 1.5. It is assumed above that the theory with relation symbols $\mathscr{R}$ and axioms A has S as its set of constants. Prove this.

Definition 1.6. Let *S be any model of $\mathscr{T}(\mathrm{S})$. We say that *S is a *standard model* of $\mathscr{T}(\mathrm{S})$ if *S is isomorphic to S, and otherwise *S is called a *non-standard model* of $\mathscr{T}(\mathrm{S})$.

Let *S = (*S, φ, ψ) be any model of $\mathscr{T}(\mathrm{S})$. Then φ: S → *S embeds S in *S, since if a, b are distinct elements of S, then $(a \neq b) \in A$ and so is true in *S, i.e., $\varphi(a) \neq \varphi(b)$. Similarly, for any n-ary relation $\rho \in \mathscr{R}$, the restriction to $\varphi(\mathrm{S})$ of the relation $\psi(\rho)$ is precisely the relation on $\varphi(\mathrm{S})$ which corresponds under φ to the relation ρ on S. We shall therefore always identify S with its image under φ in *S, and so regard the model *S as containing the standard model S.

§2 Reduction of the Order

First-order logic does not permit us to study properties of relations, or to discuss statements such as "For all n-ary relations, . . .". This restriction excludes from consideration most of the material in a subject such as real analysis, where functions of various types occupy a dominant place. The general consideration of properties of relations requires a higher-order logic. Fortunately, there is a trick which enables us to bring within the scope of our first-order predicate calculus all these higher-order concepts for any one mathematical system. For any set S, let rel(S) denote the set of all relations on S.

Definition 2.1. Let S be any set. We define the set $\mathscr{O}^k(S)$, of *kth-order objects* on S, by $\mathscr{O}^0(S) = S$, and $\mathscr{O}^{k+1}(S) = \mathscr{O}^k(S) \cup \mathrm{rel}(\mathscr{O}^k(S))$. Further, we put $\mathscr{O}^\infty(S) = \bigcup_{k \geqslant 0} \mathscr{O}^k(S)$.

For each n, we introduce an $(n + 1)$-ary relation symbol $\in^n$. If ρ is an n-ary relation on S, and if $a_1, \ldots, a_n \in S$, we can now formalise the statement that $(a_1, \ldots, a_n)$ is in ρ as $\in^n(\rho, a_1, \ldots, a_n)$, as well as by $\rho(a_1, \ldots, a_n)$. We have made ρ into an individual constant of a larger theory, and we may if we wish omit ρ from the set of relation symbols. Among the unary relations on S, there is S itself, and those elements of the extended system $\mathcal{O}^\infty(S)$ which belong to S are distinguished as those for which the formal statement $\in^1(S, a)$ is true. The statement that $\rho \in \mathcal{O}^1(S)$ is an n-ary relation on S can be formalised as

$$\tau_n(\rho) = (\exists x_1) \cdots (\exists x_n)(\in^1(S, x_1) \wedge \cdots \wedge \in^1(S, x_n) \wedge \in^n(\rho, x_1, \ldots, x_n)),$$

while the statement that ρ is an n-ary relation on the subsets $S_1, \ldots, S_n$ of S can be formalised as

$$\begin{aligned}\tau_n(\rho, S_1, \ldots, S_n) = \tau_n(\rho) \wedge (\forall x_1) \cdots (\forall x_n)(\in^n(\rho, x_1, \ldots, x_n)\\ \Rightarrow \in^1(S_1, x_1) \wedge \cdots \wedge \in^1(S_n, x_n)).\end{aligned}$$

We can now handle second-order concepts on a standard system S by forming $\mathcal{T}(\mathcal{O}^1(\mathrm{S}))$, where the set of relation symbols includes the symbols $\in^n$ and those required for the properties of relations we wish to study. The statement that all n-ary relations have the property π can then be formalised as $(\forall x)(\tau_n(x) \Rightarrow \pi(x))$.

This process may be applied to still higher-order concepts. $(k + 1)$th-order objects can be studied in $\mathcal{T}(\mathcal{O}^k(\mathrm{S}))$, which we call the $(k + 1)$th-order theory of S. We call $\mathcal{O}^k(\mathrm{S})$, together with an appropriate set of relation symbols, a *$(k + 1)$th-order standard (mathematical) system*, and $\mathcal{O}^\infty(\mathrm{S})$ an *infinite order standard system*. We point out that $\mathcal{T}(\mathcal{O}^k(\mathrm{S}))$ (including the case $k = \infty$) is still a first-order theory of an elementary standard system, namely the system $\mathcal{O}^k(\mathrm{S})$. Theorems proved about elementary standard systems thus become applicable to higher-order standard systems.

§3 Enlargements

We recall the definition of a definable n-ary relation ρ on a standard system S. We say that ρ is definable in $\mathcal{T}(\mathrm{S})$ if there is an element $p(x_1, \ldots, x_n) \in P(\mathrm{S})$, where $x_1, \ldots, x_n \in V - \mathrm{S}$ and $\mathrm{var}(p(x_1, \ldots, x_n)) \subseteq \{x_1, \ldots, x_n\} \cup \mathrm{S}$, such that

$$\rho = \{(a_1, \ldots, a_n) \in \mathrm{S}^n | p(a_1, \ldots, a_n) \text{ is true in S}\}.$$

Any such $p(x_1, \ldots, x_n)$ is called a *description* of ρ. In our work, it will not matter which description of a definable relation ρ we use. We write $\rho(x_1, \ldots, x_n)$ for some description of ρ. In the special case where ρ is a definable subset of S, we use $(x \in \rho)$ to denote an arbitrary description of ρ.

Let S be a standard system, and let $\rho(x, y)$ define a binary relation in $\mathcal{T}(\mathrm{S})$. We define the *domain* of ρ to be the set D_ρ, where

$$D_\rho = \{a \in \mathrm{S} | \rho(a, b) \text{ is true in S for some } b \in \mathrm{S}\}.$$

Definition 3.1. A *concurrent relation* of S is a definable binary relation $\rho = \rho(x, y)$ in $\mathcal{T}(\mathrm{S})$ such that $D_\rho \neq \varnothing$ and, for every finite subset $\{a_1, \ldots, a_n\}$ of D_ρ, there is a $b \in \mathrm{S}$ such that $\rho(a_i, b)$ is true for $i = 1, \ldots, n$.

Example 3.2. We consider the (elementary) real number system **R**. In **R**, $<$ is a concurrent relation with domain $D_< = \mathbf{R}$. For any finite set $\{x_1, \dots, x_n\}$ of real numbers, $y = 1 + \max_i(x_i)$ satisfies $x_i < y$ for $i = 1, \dots, n$.

Now let σ be any set of concurrent relations of the standard system S. For each $\rho \in \sigma$, we take a new variable $c_\rho \notin V$ and form $V^\sigma = V \cup \{c_\rho | \rho \in \sigma\}$, $P^\sigma = P(V^\sigma, \mathscr{R})$. We put

$$A_\sigma = \{\rho(x, c_\rho) | \rho \in \sigma \text{ and } x \in D_\rho\}.$$

Definition 3.3. The *enlargement* of $\mathscr{T} = (\mathscr{R}, A, \mathrm{S})$ with respect to σ is the theory $\mathscr{T}^\sigma = (\mathscr{R}, A \cup A_\sigma, \mathrm{S} \cup \{c_\rho | \rho \in \sigma\})$. When σ is the set of all concurrent relations of S, we call $\mathscr{T}^\sigma$ the *full enlargement* of $\mathscr{T}$, and denote it by $^*\mathscr{T}$.

Theorem 3.4. *Let $\mathscr{T} = \mathscr{T}(\mathrm{S})$ and let σ be any set of concurrent relations of* S. *Then $\mathscr{T}^\sigma$ is consistent.*

Proof. Suppose $\mathscr{T}^\sigma \vdash F$. Then $A_0 \vdash_{\mathscr{T}} F$ for some finite subset A_0 of A_σ. Let $A_0 = \{\rho_j(x_{ij}, c_{\rho_j}) | i = 1, \dots, r_j;\ j = 1, \dots, n\}$, where $x_{ij} \in D_{\rho_j}$. Since $\{x_{1j}, \dots, x_{r_j j}\}$ is a finite subset of D_{ρ_j}, there exists $b_j \in \mathrm{S}$ such that $\rho_j(x_{ij}, b_j)$ is true in S for $i = 1, \dots, r_j$. Mapping c_{ρ_j} to b_j for $j = 1, \dots, n$ makes S a model of the theory $\mathscr{T}' = (\mathscr{R}, A \cup A_0, \mathrm{S} \cup \{c_{\rho_j} | j = 1, \dots, n\})$. Hence $\mathscr{T}'$ is consistent, which contradicts the assumption that $A \cup A_0 \vdash F$. Thus $\mathscr{T}^\sigma$ is consistent. □

Since $\mathscr{T}^\sigma$ is consistent, it has a model. Let S^σ be any model of $\mathscr{T}^\sigma$. As we have already indicated, S^σ has the standard model S of $\mathscr{T}$ embedded in it. We call S^σ a σ-enlargement of S. A model $^*\mathrm{S}$ of $^*\mathscr{T}$ is called a full enlargement of S.

Suppose that $\mathrm{S}^\sigma = \mathrm{S}$. Then for each $\rho \in \sigma$, the constant c_ρ of $\mathscr{T}^\sigma$ is interpreted as some $b_\rho \in \mathrm{S}$ which satisfies $\rho(x, b_\rho)$ for all $x \in D_\rho$. Thus all we have achieved is the introduction of a new name c_ρ for the element b_ρ. The new axioms $\rho(x, c_\rho)$ reduce to axioms of $\mathscr{T}$ if we replace c_ρ by b_ρ, and so if we add to $\mathscr{T}^\sigma$ the further axioms $c_\rho = b_\rho$ for all $\rho \in \sigma$, the resulting theory is equivalent to $\mathscr{T}$ in the sense that the two theories have the same models. Such an enlargement $\mathscr{T}^\sigma$ is of little use in studying $\mathscr{T}$ and is called a trivial enlargement.

Exercise 3.5. Use the ultrapower construction studied in §5 of Chapter VII to give an alternative proof of Theorem 3.4.

One standard system may be contained in another, as in the case of the integers **Z** and the reals **R**. We shall now obtain a useful result on enlargements of systems related in this way.

Definition 3.6. Let $\mathrm{S} = (S, \mathscr{R})$ and $\mathrm{S}_1 = (S_1, \mathscr{R}_1)$ be standard systems. We say that S is a *subsystem* of S_1, and write $\mathrm{S} \leqslant \mathrm{S}_1$, if $\mathscr{R} \subseteq \mathscr{R}_1$ and if S is a definable subset of S_1.

Examples

3.7. For any S, $S \leqslant \mathcal{O}^1(S) \leqslant \mathcal{O}^2(S) \leqslant \cdots \leqslant \mathcal{O}^\infty(S)$.

3.8. Take $\mathbb{R} = (R, \{\mathscr{I}, +, \times, <, i\})$, where $i(x)$ is interpreted as "x is an integer". Then $\mathbb{Z} = (Z, \{\mathscr{I}, +, \times, <\})$ is a subsystem of $\mathbb{R}$.

3.9. $\mathcal{O}^k(\mathbb{Z}) \leqslant \mathcal{O}^{k+1}(\mathbb{R})$ for all $k \geqslant 0$.

Let $S \leqslant S_1$, and let $\rho(x, y)$ be a concurrent relation of S. Since S is definable, the relation ρ on S can be defined in $P(S_1)$ by

$$\rho_1(x, y) = (x \in S) \wedge (y \in S) \wedge \rho(x, y).$$

This ρ_1 is a concurrent relation on S_1, consisting of precisely those pairs of elements which are in ρ. In general, if σ, σ_1 are sets of concurrent relations of S, S_1 respectively, we say $\sigma \leqslant \sigma_1$ if, for every $\rho \in \sigma$, we have that the corresponding $\rho_1 \in \sigma_1$.

Theorem 3.10. *Let* $S \leqslant S_1$, *and let* σ, σ_1 *be sets of concurrent relations of* S, S_1 *such that* $\sigma \leqslant \sigma_1$. *Let* $S_1^{\sigma_1}$ *be an enlargement of* S_1 *with respect to* σ_1. *Then* $S_1^{\sigma_1}$ *contains an enlargement* S^σ *of* S *with respect to* σ. *In particular, any full enlargement of* S_1 *contains a full enlargement of* S.

Proof. Put $S^\sigma = \{a \in S_1^{\sigma_1} | p(a) \text{ is true in } S_1^{\sigma_1}\}$, where $p(x) \in P(S_1)$ is such that $S = \{a \in S_1 | p(a) \text{ is true in } S_1\}$. S^σ is clearly a model of $\mathscr{T}(S)$, and we must show that for each $\rho \in \sigma$, there is a $b_\rho \in S^\sigma$ such that $\rho(a, b_\rho)$ is true in S^σ for all $a \in D_\rho$. Now ρ_1 is a concurrent relation of S_1, hence there is a $b_{\rho_1} \in S_1^{\sigma_1}$ such that $\rho_1(a, b_{\rho_1})$ is true in $S_1^{\sigma_1}$ for all $a \in D_{\rho_1}$. But $D_{\rho_1} = D_\rho$, and by the construction of ρ_1, $\rho_1(a, b_{\rho_1})$ true implies that $b_{\rho_1} \in S^\sigma$. Thus we can take $b_\rho = b_{\rho_1}$. □

§4 Standard Relations

Definition 4.1. Let S^σ be an enlargement of S. A *standard n-ary relation* on S^σ is a relation

$$\rho^\sigma = \{(a_1, \ldots, a_n) \in (S^\sigma)^n | \rho(a_1, \ldots, a_n) \text{ is true in } S^\sigma\}$$

for some definable n-ary relation ρ on S.

We also define a *standard element* of S^σ to be an element of S.

Exercises

4.2. Show that the standard relation ρ^σ is independent of the choice of description of ρ.

4.3. Show that the one-element subset $\{a\}$ of S^σ is standard if and only if a is standard.

Theorem 4.4. *Let* *S *be a full enlargement of* S *and let* u *be a definable subset of* S. *Then* $^*u = u$ *if and only if* u *is finite.*

Proof. Suppose that $u = \{u_1, \ldots, u_n\}$ is finite. Then

$$u(x) = (x = u_1) \vee (x = u_2) \vee \cdots \vee (x = u_n)$$

is a description of u, and

$$^*u = \{a \in {}^*\mathrm{S} \mid (a = u_1) \vee (a = u_2) \vee \cdots \vee (a = u_n) \text{ is true in } {}^*\mathrm{S}\} = u.$$

Suppose that u is infinite. Let ρ be the binary relation on S defined by $\rho(x, y) = (x \in u) \wedge (y \in u) \wedge (x \neq y)$. Then $D_\rho = u$ and, since u is infinite, for any $u_1, \ldots, u_n \in D_\rho$, there exists $y \in u$ distinct from $u_1, \ldots, u_n$ and thus satisfying $\rho(u_i, y)$ for all i. Therefore ρ is a concurrent relation, and so there is a $b_\rho \in {}^*\mathrm{S}$ such that $\rho(x, b_\rho)$ is true for all $x \in D_\rho$. This says that $b_\rho \in {}^*u$ and $b_\rho \notin u$. □

Corollary 4.5. *Suppose the enlargement $\mathscr{T}^\sigma$ of $\mathscr{T}(\mathrm{S})$ is both full and trivial. Then* S *is finite.*

Proof. S is a definable subset with description $p(x) = {\sim}F$. □

Corollary 4.6. *Let ρ be a definable n-ary relation on* S. *Then $^*\rho = \rho$ if and only if ρ is finite.*

Proof. If ρ is finite, we can give a description which lists its members and it follows that $^*\rho = \rho$. If ρ is infinite, put

$$u_i(x) = (\exists x_1) \cdots (\exists x_{i-1})(\exists x_{i+1}) \cdots (\exists x_n)\rho(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n),$$

and let u_i be the subset of S defined by $u_i(x)$. Then for some i, u_i is infinite, and the theorem implies that $^*u_i \neq u_i$, i.e., that $^*\rho \neq \rho$. □

Theorem 4.7. *Let ρ be a definable relation on* S *which defines a function $f: D \to \mathrm{S}$ on some definable*[2] *subset D of* S. *Then ρ^σ defines a function $f^\sigma: D^\sigma \to \mathrm{S}^\sigma$ on the subset D^σ of S^σ.*

Proof. We have $\mathscr{T} \vdash (\forall x)((x \in D) \Rightarrow (\exists! y)\rho(x, y))$. Interpreting this in S^σ gives the result. □

The same argument applies to show that if $f: U \to V$ is a definable function, where U is a definable subset of S^r and V is a definable subset of S^s, then f^σ is a function from U^σ to V^σ.

§5 Internal Relations

Let S be any standard system. For each n, let $\mathscr{R}_n^{(1)}$ be the set of all first-order n-ary relations on S. Then each element of $\mathscr{R}_n^{(1)}$, and also the set $\mathscr{R}_n^{(1)}$ itself are all definable in $\mathscr{T}(\mathscr{O}^1(\mathrm{S}))$. If $(\mathscr{O}^1(\mathrm{S}))^\sigma$ is an enlargement of $\mathscr{O}^1(\mathrm{S})$ with respect to some set σ of concurrent binary relations of $\mathscr{O}^1(\mathrm{S})$, then every element of $(\mathscr{R}_n^{(1)})^\sigma$ is an n-ary relation on S^σ.

[2] Note that $\{a \in \mathrm{S} \mid (\exists! y)\rho(a, y) \text{ is true}\}$ is a possible choice for D, but it is not always the most convenient choice.

Definition 5.1. An *internal* first-order n-ary relation on S^σ is an element of $(\mathscr{R}_n^{(1)})^\sigma$. A relation on S^σ which is not internal is called *external.*

Higher-order internal n-ary relations may be defined similarly, by using the set $\mathscr{R}_n^{(k)}$ of n-ary relations on $\mathscr{O}^{k-1}(S)$.

If $\rho \in \mathscr{R}_n^{(1)}$ is in fact a definable relation on S, we have that

$$\mathscr{T}(\mathscr{O}^1(S)) \vdash (\forall x_1)\cdots(\forall x_n)(\in^n(\rho, x_1, \ldots, x_n) \Rightarrow \rho(x_1, \ldots, x_n)),$$

while each $u \in (\mathscr{R}_n^{(1)})^\sigma$ is the relation

$$u = \{(a_1, \ldots, a_n) \in (\mathscr{O}^1(S))^\sigma | \in^n(u, a_1, \ldots, a_n) \text{ is true in } (\mathscr{O}^1(S))^\sigma\}.$$

It follows that the definable relation ρ, considered as an element of $(\mathscr{R}_n^{(1)})^\sigma$, is the relation ρ^σ on S^σ. Hence every standard relation is internal. The converse is not true, for if S is infinite, then $\mathscr{R}_n^{(1)}$ is infinite, and by Theorem 4.4, $^*(\mathscr{R}_n^{(1)}) \neq \mathscr{R}_n^{(1)}$.

Lemma 5.2. *Let u, v be internal n-ary relations on S^σ. Then $u \cap v$, $u \cup v$, and the complement $u^\sim$ of u are also internal.*

Proof. We have that

$$\mathscr{T}(\mathscr{O}^1(S)) \vdash (\forall x)(\forall y)((x \in \mathscr{R}_n^{(1)}) \wedge (y \in \mathscr{R}_n^{(1)}) \Rightarrow$$
$$(\exists z)((z \in \mathscr{R}_n^{(1)}) \wedge (\forall t)(t \in z \Leftrightarrow (t \in x) \wedge (t \in y)))).$$

It follows that $u \cap v \in (\mathscr{R}_n^{(1)})^\sigma$ for all $u, v \in (\mathscr{R}_n^{(1)})^\sigma$. Similar proofs apply for $u \cup v$ and for $u^\sim$. □

§6 Non-Standard Analysis

Let **R** be the set of real numbers, with relation symbols $\mathscr{R} = \{\mathscr{I}, \times, +, <\}$. We form a full enlargement $^*(\mathscr{O}^k(\mathbf{R}))$ for some $k \geqslant 1$. Within this, we have standard subsets $^*\mathbf{R} > {}^*\mathbf{Q} > {}^*\mathbf{Z} > {}^*\mathbf{N}$, which are full enlargements of the reals, rationals, integers and natural numbers respectively. The relations on $^*\mathbf{R}$ defined by $\times, +, <$ shall be denoted by the same symbols, instead of by the correct but more cumbersome $^*\times$, etc. The function $|\,|: \mathbf{R} \to \mathbf{R}$, defined by $|x| = x$ if $x \geqslant 0$, $|x| = -x$ if $x < 0$, yields the standard function $^*\mathbf{R} \to {}^*\mathbf{R}$ defined in the same way and which we shall denote by the same symbol $|\,|$. We shall call the elements of $^*\mathbf{R}$ real numbers, distinguishing those in **R** by calling them standard real numbers.

Theorem 6.1. *$^*\mathbf{R}$ is a non-archimedean ordered field.*

Proof. The axioms of ordered fields are theorems of $\mathscr{T}(\mathbf{R})$ and so hold for $^*\mathbf{R}$, showing that $^*\mathbf{R}$ is an ordered field. The relation $x < y$ is a concurrent relation of **R** with domain **R**, and consequently there is an element $a \in {}^*\mathbf{R}$ such that $r < a$ for all $r \in \mathbf{R}$. This implies that for any $r \in \mathbf{R}$, and for all $n \in \mathbf{N}$, $\sum_{i=1}^{n} r < a$. Hence the ordering on $^*\mathbf{R}$ is non-archimedean. □

The archimedean axiom can indeed be expressed in the language of $\mathcal{O}^k(\mathrm{R})$ as

$$(\forall x)(\forall y)((x \in \mathrm{R}) \wedge (y \in \mathrm{R}) \wedge (x > 0) \Rightarrow (\exists n)((n \in \mathrm{N}) \wedge (nx > y))),$$

where $nx > y$ is an abbreviation for $(\forall z)(\times(n, x, z) \Rightarrow (z > y))$. This is a theorem of $\mathscr{T}(\mathcal{O}^k(\mathrm{R}))$ and so holds in $^*\mathrm{R}$. It does not assert the archimedean property for $^*\mathrm{R}$, as it asserts that if $x, y \in {}^*\mathrm{R}$ and if $x > 0$, then there is $n \in {}^*\mathrm{N}$ such that $nx > y$. For $^*\mathrm{R}$ to be archimedean, we need to have $n \in \mathrm{N}$.

Definition 6.2. An element $a \in {}^*\mathrm{R}$ is called *finite* if there exists a standard real number b such that $|a| < b$. Otherwise, a is called *infinite*. A (finite) element a is called *infinitesimal* if $|a| < b$ for all standard real numbers $b > 0$.

0 is infinitesimal, and since $0 < a < b$ holds if and only if $0 < 1/b < 1/a$, it follows that if $a \neq 0$, then a is infinitesimal if and only if $1/a$ is infinite.

The proof of Theorem 6.1 contains a proof of the existence of infinite real numbers, and it follows that infinite natural numbers also exist.

Lemma 6.3. *There is no smallest infinite natural number. The set of infinite natural numbers is an external set.*

Proof. If n is a natural number and $n \neq 0$, then $n = m + 1$ for some natural number m, since this result is a theorem of $\mathscr{T}(\mathrm{N})$. If n is the smallest infinite natural number, then $m = n - 1$ is also infinite, and $m < n$, giving a contradiction.

It is a theorem of $\mathscr{T}(\mathcal{O}^1(\mathrm{N}))$ that every non-empty subset of N has a least member. Hence every non-empty internal subset of $^*\mathrm{N}$ has a least member, and the set of infinite natural numbers cannot be internal. □

Lemma 6.4. *Suppose $n \in {}^*\mathrm{N}$. Then n is finite if and only if $n \in \mathrm{N}$.*

Proof. If $n \in \mathrm{N}$, n is clearly finite. Suppose that n is finite. Then $n < b$ for some standard real number b, and $b < m$ for some standard natural number m. Put $u = \{x \in \mathrm{N} | x < m\}$. Then $n \in {}^*u = \{x \in {}^*\mathrm{N} | x < m\}$, and $^*u = u$ since u is finite. Thus $n \in \mathrm{N}$. □

Theorem 6.5. *Each of* N, R, *the set of infinite real numbers, and the set of infinitesimal real numbers is an external set.*

Proof.

(a) By Lemma 6.3, the set of infinite natural numbers is an external set, and by Lemma 6.4, N is its complement in the internal set $^*\mathrm{N}$. Hence N is external by Lemma 5.2.

(b) If R is internal, then so is $\mathrm{N} = \mathrm{R} \cap {}^*\mathrm{N}$, contradicting (a). Similarly, Z and Q are also external.

(c) Let R_∞ be the set of infinite real numbers. If it is internal, then so is $R_\infty \cap {}^*\mathrm{N}$, contradicting Lemma 6.3.

(d) If R_1 is the set of infinitesimal real numbers, then R_1 is bounded above and has no greatest member. It is a theorem of $\mathscr{T}(\mathscr{O}^1(\mathrm{R}))$ that if u is a non-empty subset of R which is bounded above and has no greatest element, then $\{x \in \mathrm{R} | x > y \text{ for all } y \in u\}$ has a least element. If R_1 is internal, then $v = \{x \in {}^*\mathrm{R} | x > r \text{ for all } r \in R_1\}$ has a least element. But if $x \in v$, then $\frac{1}{2}x \in v$ and $\frac{1}{2}x < x$. Hence v has no least element, and so R_1 is external. □

Let $a, b \in {}^*\mathrm{R}$. We write $a \doteq b$ if $a - b$ is infinitesimal. $\doteq$ is clearly an equivalence relation on *R.

Exercise 6.6. Show that if $r \in {}^*\mathrm{R}$, then there exists $q \in {}^*\mathrm{Q}$ such that $q \doteq r$.

Definition 6.7. The *monad* of the finite real number a is the set $\mu(a) = \{r \in {}^*\mathrm{R} | r \doteq a\}$.

Theorem 6.8. *If a is a finite real number, then $\mu(a)$ contains exactly one standard real number. If R_0 is the set of finite real numbers and R_1 the set of infinitesimal real numbers, then R_0 is a ring, R_1 is an ideal of R_0 and R_0/R_1 is isomorphic to* R.

Proof. If $r, s \in \mu(a)$ and r, s are standard, then $|r - s|$ is an infinitesimal standard real number. Thus $|r - s| = 0$ and $r = s$. We have to show that there is a standard real number in $\mu(a)$. This is so if a is standard, so we suppose a is not standard. Put $L = \{x \in \mathrm{R} | x < a\}$ and $U = \{x \in \mathrm{R} | x > a\}$. Since a is finite, there exists $b \in \mathrm{R}$ such that $|a| < b$, i.e., $-b < a < b$, showing that L and U are both non-empty. L is bounded above by b and so has a least upper bound α say, which is also the greatest lower bound of U. If $\alpha \in L$, then $U = \{x \in \mathrm{R} | x > \alpha\}$, and $\alpha \leqslant a < \alpha + r$ for all standard real numbers $r > 0$. Thus $|a - \alpha| = a - \alpha < r$ for all standard $r > 0$, and so $a - \alpha$ is infinitesimal. Similarly, if $\alpha \in U$, we obtain $|a - \alpha| = \alpha - a$ is infinitesimal. Hence $\alpha \in \mu(a)$.

Trivially, R_0 is a ring and R_1 is an ideal of R_0. The map sending a to $\mu(a)$ is the natural homomorphism $R_0 \to R_0/R_1$. Mapping $\mu(a)$ to the standard real number in $\mu(a)$ is an isomorphism. □

Finally, as an introduction to the use of enlargements in the study of analysis, we shall show how a few of the familiar results on limits can be proved with the aid of infinitesimal and infinite elements in an enlargement. We begin with the concept of a limit of a sequence. A sequence is a function $s\colon \mathrm{N} \to \mathrm{R}$, and corresponding to any sequence, we have the standard function ${}^*s\colon {}^*\mathrm{N} \to {}^*\mathrm{R}$.

Theorem 6.9. *Let $r \in \mathrm{R}$ and let $s\colon \mathrm{N} \to \mathrm{R}$ be a sequence. Then* $\mathrm{Lim}_{n \to \infty} s(n) = r$ *if and only if ${}^*s(n) \in \mu(r)$ for all infinite natural numbers n.*

Proof. Suppose that $\mathrm{Lim}_{n \to \infty} s(n) = r$. Then for every standard real number $\varepsilon > 0$, there exists $n_0 \in \mathrm{N}$ such that $|s(n) - r| < \varepsilon$ for all $n > n_0$. For this ε and n_0, $(\forall n)((n \in \mathrm{N}) \wedge (n > n_0) \Rightarrow |s(n) - r| < \varepsilon)$ is a theorem of

$\mathscr{T}(\mathscr{O}^1(\mathrm{R}))$. Therefore $(\forall n)((n \in {}^*\mathrm{N}) \wedge (n > n_0) \Rightarrow |{}^*s(n) - r| < \varepsilon)$ holds in ${}^*(\mathscr{O}^1(\mathrm{R}))$. If n is an infinite natural number, then $|{}^*s(n) - r| < \varepsilon$, and this is true for all standard real numbers $\varepsilon > 0$. Hence ${}^*s(n) \in \mu(r)$ for all infinite numbers n.

Suppose conversely that ${}^*s(n) \in \mu(r)$ for all infinite natural numbers n. If n_0 is an infinite natural number, then for every standard real number $\varepsilon > 0$, we have $|{}^*s(n) - r| < \varepsilon$ for all $n > n_0$. Thus $(\exists n_0)((n_0 \in {}^*\mathrm{N}) \wedge (\forall n)((n \in {}^*\mathrm{N}) \wedge (n > n_0) \Rightarrow |{}^*s(n) - r| < \varepsilon))$, being true in ${}^*\mathscr{O}^1(\mathrm{R})$, is a theorem of $\mathscr{T}(\mathscr{O}^1(\mathrm{R}))$. Hence there exists $n_0 \in \mathrm{N}$ such that $|s(n) - r| < \varepsilon$ for all $n > n_0$. □

By a similar argument, one obtains the following result.

Theorem 6.10. *Let U be a subset of* R, *and suppose U contains a neighbourhood of $a \in \mathrm{R}$. Let $f: U \to \mathrm{R}$ be a function defined on U. If $\ell \in \mathrm{R}$, then $Lim_{x \to a} f(x) = \ell$ if and only if ${}^*f(x) \in \mu(\ell)$ for all $x \neq a$ in $\mu(a)$.*

Corollary 6.11. *The function f is continuous at a if and only if ${}^*f(x) \simeq {}^*f(a)$* for all $x \simeq a$.

Exercise 6.12. Prove Theorem 6.10.

For a real function f defined on an arbitrary subset U of R, the necessary and sufficient condition that f be continuous on U is that for each $a \in U$, if $x \in {}^*U$ and $x \simeq a$, then ${}^*f(x) \simeq {}^*f(a)$. The meaning of this condition is altered if we write it formally using $(\forall a)$, as we now show. For then the statement becomes the following: for all $a, x \in {}^*U$, if $x \simeq a$ then ${}^*f(x) \simeq {}^*f(a)$. If this new statement holds, then for an infinitesimal positive real number δ, and for any standard real number $\varepsilon > 0$,

$$(\forall a)(\forall x)((a \in {}^*U) \wedge (x \in {}^*U) \wedge (|x - a| < \delta) \Rightarrow |{}^*f(x) - {}^*f(a)| < \varepsilon)$$

is true in ${}^*(\mathscr{O}^1(\mathrm{R}))$, and so

$$(\exists \delta)((\delta > 0) \wedge (\forall a)(\forall x)((a \in U) \wedge (x \in U) \\ \wedge (|x - a| < \delta) \Rightarrow |f(x) - f(a)| < \varepsilon))$$

holds in $\mathscr{O}^1(\mathrm{R})$. But this is precisely the condition that the function f be *uniformly* continuous on U. We have proved the following theorem.

Theorem 6.13. *Let f be a real-valued function defined on the subset U of* R. *Then f is uniformly continuous on U if and only if for all $x, y \in {}^*U$, $x \simeq y$ implies that ${}^*f(x) \simeq {}^*f(y)$.*

It is now a simple matter to prove the following well-known result.

Theorem 6.14. *Let U be a closed bounded interval $[p, q]$. If the real-valued function f is continuous on U, then it is uniformly continuous on U.*

Proof. Take any $x \in {}^*U$. x is a finite real number, hence there is a unique $r \in \mathrm{R}$ such that $r \simeq x$. If $r < p$, then $x = r + (x - r) < r + (p - r) = p$, since $x - r$ is infinitesimal and $p - r$ is a standard positive real number. As $x \geqslant p$, we have a contradiction, and so $p \leqslant r$. Similarly, $r \leqslant q$, and $r \in U$. If $y \in {}^*U$ and $y \simeq x$, then $y \simeq r$ and ${}^*f(y) \simeq {}^*f(r)$. In particular, ${}^*f(x) \simeq {}^*f(r)$.

Consequently $^*f(x) \simeq {}^*f(r) \simeq {}^*f(y)$, and we have the condition for uniform continuity on U. □

Exercise 6.15. Where does the above method of proof fail if U is taken as the open interval $\{x : p < x < q\}$?

Our final application is to the study of sequences of real-valued functions $s_n(x)$ defined on a subset U of **R**. The usual necessary and sufficient condition that $s_n(x) \to r(x)$ on U as $n \to \infty$, when expressed in terms of our non-standard analysis, is that for each $x \in U$ and for all infinite n, $^*s_n(x) \simeq {}^*r(x)$. Again, the meaning of the condition is altered if we express it in terms of $(\forall x)$, as the next result indicates.

Theorem 6.16. *The sequence of functions $s_n(x)$ converges uniformly on U to $r(x)$ if and only if for all $x \in {}^*U$ and for all infinite n, $^*s_n(x) \simeq {}^*r(x)$.*

Exercise 6.17. Prove Theorem 6.16 by suitably modifying the argument leading to Theorem 6.13.

Theorem 6.18. *Suppose the functions $s_n(x)$ are continuous on U, and converge uniformly on U to $r(x)$. Then $r(x)$ is continuous on U.*

Proof. Let $a \in U$. If $x \in {}^*U$ and $x \simeq a$, and if n is infinite, then by Corollary 6.11 and Theorem 6.14,

$$^*r(s) \simeq {}^*s_n(x) \simeq {}^*s_n(a) \simeq {}^*r(a);$$

showing that r is continuous at a. □

Exercises

6.19. Given that $f(x) \to r$ and $g(x) \to s$ $(\neq 0)$ as $x \to a$, prove that $f(x)/g(x) \to r/s$ as $x \to a$.

6.20. $f(x)$ is defined in a neighbourhood of a. Show that $f'(a) = c$ if and only if $\dfrac{^*f(x) - {}^*f(a)}{x - a} \simeq c$ for all $x \simeq a$, $x \neq a$.

6.21. Prove that if $f(x)$ is differentiable at $x = a$, then $f(x)$ is continuous at $x = a$.

6.22. **R** is complete, i.e., every Cauchy sequence in **R** has a limit in **R**. Formalise this and interpret it for ***R**. Is ***R** complete?

We refer the reader to [8] for further reading and references on the subject of non-standard analysis.

Chapter IX

Turing Machines and Gödel Numbers

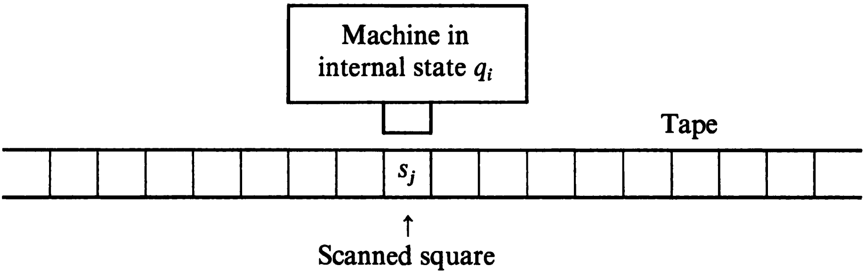
§1 Decision Processes

In §3 of Chapter III, we gave a procedure for determining whether or not an element p of $P(X)$ is a theorem of Prop(X). In §4 of Chapter IV, we asserted that no such procedure exists for Pred$(V, \mathscr{R})$. Before attempting to prove this non-existence theorem, we must say more precisely what we mean by "procedure". The procedures we shall discuss are called decision processes, and informally we think of a decision process as a list of instructions which can be applied in a routine fashion to give one of a finite number of specified answers. A decision process for Pred$(V, \mathscr{R})$ is then a finite list of instructions such that for any element $p \in P(V, \mathscr{R})$, there corresponds a unique finite sequence of instructions from the list. The sequence terminates with an instruction to announce a decision of some prescribed kind (e.g., "p is a theorem of Pred$(V, \mathscr{R})$."). Thus at each step of the process, exactly one instruction of the list is applicable, producing a result to which exactly one instruction is applicable, until after a finite (but not necessarily bounded) number of steps, the process stops and a decision is announced.

The mechanical nature of the process just described suggests that we could think of it as a computer program, carried out on a suitable computer. We shall formalise our ideas by considering processes which could be performed by an idealised computer known as a Turing machine.

§2 Turing Machines

A Turing machine is imagined as consisting of two parts—the machine proper, being a device with a finite set $\mathfrak{Q} = \{q_0, q_1, \ldots, q_m\}$ of possible internal states, and a tape (at least potentially infinite) on which suitably coded instructions to the machine may be printed, and on which the machine can print its response. The tape is divided lengthwise into squares which can be indexed by the integers $\mathbf{Z}$. On each square of the tape is printed one symbol selected from a fixed finite set $\mathfrak{S} = \{s_0, s_1, \ldots, s_k\}$, called the alphabet of the machine. Since we think in terms of finite lists of instructions, we must allow squares to be blank, and hence the alphabet $\mathfrak{S}$ must contain a symbol corresponding to 'blank'. This symbol will always be denoted by s_0. Only finitely many squares of the tape have printed on them a symbol other than s_0. The tape is fed into the machine so that at any time, the machine is scanning exactly one square of the tape.



We assume that the machine must be in internal state q_0 to commence operating. The machine operates in discrete steps, its action at any stage being determined by its internal state q_i together with the symbol s_j printed on the square being scanned. The possible actions of the machine are of the following kinds:

(i) The machine replaces the symbol s_j by a symbol s_ℓ and changes its internal state to q_r.

(ii) The machine moves the tape so as to scan the square immediately on the right of the one being scanned, and changes its internal state to q_r.

(iii) The machine moves the tape so as to scan the square immediately on the left of the one being scanned, and changes its internal state to q_r.

(iv) The machine stops.

Since the machine must have no choice of action, exactly one of the above actions will occur at each step.

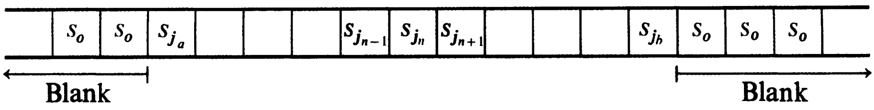
A Turing machine is specified by giving its set $\mathfrak{Q}$ of internal states, its alphabet $\mathfrak{S}$, and its response to each pair (q_i, s_j) consisting of an internal state and a scanned symbol. Since there are only finitely many pairs (q_i, s_j), the machine response is specified by a finite list. A response of the type (i) can be indicated by quadruples (q_i, s_j, s_ℓ, q_r). Responses (ii) and (iii) can be indicated by the quadruples (q_i, s_j, R, q_r) and (q_i, s_j, L, q_r) respectively, where we have assumed that neither R nor L is in $\mathfrak{S}$. Response (iv) can be specified by having no quadruple beginning with the pair q_i, s_j. Our requirement that the machine be deterministic means that the list of responses has at most one quadruple beginning with each pair q_i, s_j.

We can now expect the following formal definition of a Turing machine to make sense.

Definition 2.1. A *Turing machine* with (finite) alphabet $\mathfrak{S}$ and (finite) set $\mathfrak{Q}$ of internal states is a subset M of $\mathfrak{Q} \times \mathfrak{S} \times (\mathfrak{S} \cup \{L, \mathscr{R}\}) \times \mathfrak{Q}$ $(L, R \notin \mathfrak{S})$, such that if (a, b, c, d) and $(a, b, c', d') \in M$, then $c = c'$ and $d = d'$.

To discuss the operation of a Turing machine M, we need a convenient way of describing its state at each stage of a computation. The state of M at any stage is determined by the contents of the tape, the number of the tape square being scanned, and the internal state of the machine[1]. Denote the

[1] Thus the state of M is a description of the total machine configuration, including the internal state of M.



symbol printed on square number n by s_{j_n}. Since there are always only finitely many non-blank squares, there exist integers a, b (not unique) such that $j_n = 0$ for all $n < a$ and all $n > b$. a and b can always be selected so that the square currently being scanned, say square number n, lies between square number a and square number b, so that $a \leqslant n \leqslant b$. (We note that only the ordering of the tape squares is important—it is customary to shift the origin each time the machine shifts the tape, so that the square being scanned becomes the origin.) The contents of the tape are completely specified by the finite string $s_{j_a} s_{j_{a+1}} \cdots s_{j_n} \cdots s_{j_b}$, and we shall indicate that the machine is in internal state q_i, scanning square n of a tape with these symbols on it, by writing the string

$$s_{j_a} s_{j_{a+1}} \cdots s_{j_{n-1}} q_i s_{j_n} \cdots s_{j_b}.$$

Definition 2.2. An *instantaneous description* of a Turing machine M with alphabet $\mathfrak{S}$ and set $\mathfrak{Q}$ of internal states is a finite string $s_{\alpha_1} s_{\alpha_2} \cdots s_{\alpha_r} q s_{\beta_1} s_{\beta_2} \cdots s_{\beta_t}$, where $s_{\alpha_i}, s_{\beta_j} \in \mathfrak{S}$ and $q \in \mathfrak{Q}$.

The strings $s_{\alpha_1} \cdots s_{\alpha_r}$ and $s_{\beta_1} \cdots s_{\beta_t}$ are often denoted by single symbols such as σ, τ. An instantaneous description $d = s_{\alpha_1} s_{\alpha_2} \cdots s_{\alpha_r} q s_{\beta_1} s_{\beta_2} \cdots s_{\beta_t}$ is then written simply as $d = \sigma q \tau$. Each of σ, τ may be the empty string.

Since we are interested in the state of M, rather than in descriptions of the state of M, we need to know when two descriptions determine the same state. The previous discussion shows that the only freedom in the definition of description is in the choice of a and b. Thus two descriptions $d = \sigma q \tau$ and $d' = \sigma' q' \tau'$ describe the same state if and only if $q = q'$, σ' is obtainable from σ by adding or deleting a number of symbols s_0 on the left, and τ' is obtainable from τ by adding or deleting a number of symbols s_0 on the right. Descriptions related in this way are called equivalent, and the equivalence class containing the description d is denoted by $[d]$ and called the *state* described by d. For each state $[d]$, there is a unique description $d = \sigma q \tau$ such that the first symbol (if any) of σ, and the last symbol (if any) of τ are distinct from s_0. This description is called the *shortest description* of $[d]$.

Definition 2.3. The Turing machine M takes the state $[d]$ into the state $[d']$, written $[d] \xrightarrow{M} [d']$, if for some representatives $d = \sigma q \tau$ and $d' = \sigma' q' \tau'$, where $\tau = s_\alpha \tau_1$, either

(i) $(q, s_\alpha, s_{\alpha'}, q') \in M$ and $\sigma' = \sigma, \tau' = s_{\alpha'} \tau_1$, or
(ii) $(q, s_\alpha, R, q') \in M$ and $\sigma' = \sigma s_\alpha, \tau' = \tau_1$, or
(iii) $(q, s_\alpha, L, q') \in M$ and $\sigma = \sigma' s_\beta, \tau' = s_\beta \tau$ for some $s_\beta \in S$.

Exercise 2.4. Prove that there is at most one state $[d']$ such that $[d] \xrightarrow{M} [d']$. When $[d']$ exists, show that to each $d \in [d]$, there corresponds a

$d' \in [d']$ so that d and d' are related as in (i), (ii) or (iii) of the definition (appropriately modified if σ or τ is empty).

Definition 2.5. A state $[\sigma q \tau]$ is called *initial* if $q = q_0$. A state $[\sigma q s_\alpha \tau_1]$ is called *terminal* if there is no quadruple (q, s_α, c, d) in M.

Exercise 2.6. Show that $[d]$ is terminal if and only if there does not exist a state $[d']$ such that $[d] \overset{M}{\to} [d']$.

Definition 2.7. A *computation* by the machine M is a finite sequence $[d_0], [d_1], \ldots, [d_p]$ of states such that $[d_0]$ is initial, $[d_p]$ is terminal and $[d_i] \overset{M}{\to} [d_{i+1}]$ for $i = 0, 1, \ldots, p - 1$.

Computations are by definition finite. Given M and $[d]$, there is no guarantee that M, started in state $[d]$ and allowed to operate, will ever stop (i.e., will execute a computation).

Definition 2.8. We say that M *fails* for the input $[d_0]$ if there is no computation by M beginning with the state $[d_0]$.

For each state $[d_i]$, there is a unique $[d_{i+1}]$ such that $[d_i] \overset{M}{\to} [d_{i+1}]$. Hence failure of M for the input $[d_0]$ means that the sequence of states taken by M and beginning with $[d_0]$ is infinite—i.e., the machine never stops.

Henceforth, the state $[d]$ will be denoted simply by some description d. The context will make clear the sense in which symbols such as d, d_i are being used.

Exercises

2.9. A stereo-Turing machine M has its tape divided into two parallel tracks. The symbols on a pair of squares (one above the other) are read simultaneously. Show that there is a (mono-)Turing machine M' which will perform essentially the same computations as M.

2.10. The operator of the Turing machine M has been asked to record the output of M (i.e., the symbols printed on the tape) at the end of each computation by M. Does the operator have any problems? Show that a machine M' can be designed so as to perform essentially the same computations as M, and which in addition will place marker symbols (not in the alphabet of M) either at the furthest out points of the tape used in each computation, or alternatively at the nearest points such that the stopping position of M', and all non-blank symbols, lie between them.

2.11. A dual-Turing machine M with alphabet $\mathfrak{S}$ has two tapes which can move independently. Show that there is a Turing machine with alphabet $\mathfrak{S} \times \mathfrak{S}$ which will, when given an initial state corresponding to the pair of initial states of a computation by M, perform a computation whose terminal state corresponds to the pair of terminal states of M.

2.12. M_1 and M_2 are Turing machines with the same alphabet $\mathfrak{S}$. A computation by M_1 and M_2 consists of a computation by each of M_1 and

M_2 such that, if $\sigma q_i \tau$ is the output of M_1, then $\sigma q_0 \tau$ is the input for M_2. Show that there is a Turing machine M, whose alphabet contains $\mathfrak{S}$, such that if M is started in an initial state of a computation by M_1 and M_2 with terminal state $\sigma q_j \tau$, then M executes a computation with terminal state $\sigma q_k \tau$ for some q_k, while M fails if started in any other initial state.

2.13. $M_1, \ldots, M_n$ are Turing machines with the same alphabet. An algorithm requires that at each step, exactly one of $M_1, \ldots, M_n$ be applied to the result of the previous step. The Turing machine M, applied to the output of any step, determines which of $M_1, \ldots, M_n$ is to be applied for the next step. Show that there is a single Turing machine which can execute the algorithm and give the same ultimate output.

2.14. Most digital computers can read and write on magnetic tape. The tapes are finite, but the operator can replace them if they run out. Show that such computers can be regarded as Turing machines. In fact, the most sophisticated computers can be regarded as Turing machines. (This is not a mathematical exercise. The reader is asked to review his experience of computers and to see that the definitions given so far are broad enough to embrace the computational features of the computers he has used.)

§3 Recursive Functions

Let M be a Turing machine with alphabet $\mathfrak{S}$. We show how to use M to associate with each pair (k, ℓ) of natural numbers a subset $U_M^{(k,\ell)}$ of $\mathbf{N}^k$ and a function $\Psi_M^{(k,\ell)}: U_M^{(k,\ell)} \to \mathbf{N}^\ell$. For $(n_1, \ldots, n_k) \in \mathbf{N}^k$, put

$$\operatorname{code}(n_1, \ldots, n_k) = s_1^{n_1} s_0 s_1^{n_2} s_0 \cdots s_1^{n_{k-1}} s_0 s_1^{n_k},$$

where the notation s^n denotes a string of n consecutive symbols s. There may or may not be a computation by M whose initial state is the state $d_0 = q_0 \operatorname{code}(n_1, \ldots, n_k)$. If there is, let $d_t = \sigma q \tau$ be its (uniquely determined) terminal state. Choose a description d_t of this terminal state which has at least ℓ occurrences of s_0 in τ, and determine $(a_1, \ldots, a_\ell) \in \mathbf{N}^\ell$ by defining a_1 to be the number of times s_1 occurs in τ before the first occurrence of s_0, and a_i (for $2 \leqslant i \leqslant \ell$) to be the number of times s_1 occurs in τ between the $(i-1)$th and the ith occurrences of s_0. Let $U_M^{(k,\ell)}$ be the subset of $\mathbf{N}^k$ consisting of all $(n_1, \ldots, n_k) \in \mathbf{N}^k$ for which there exists a computation by M with initial state $q_0 \operatorname{code}(n_1, \ldots, n_k)$, and so for which an element $(a_1, \ldots, a_\ell) \in \mathbf{N}^\ell$ is defined. The function $\Psi_M^{(k,\ell)}$, with domain $U_M^{(k,\ell)}$, is defined by the rule

$$\Psi_M^{(k,\ell)}(n_1, \ldots, n_k) = (a_1, \ldots, a_\ell).$$

Definition 3.1. A function $\Psi_M^{(k,\ell)}$ defined as above in terms of a Turing machine M is called a *partial recursive function*[2]. The function $\Psi_M^{(k,\ell)}$ is called a (*total*) *recursive function* if $U_M^{(k,\ell)} = \mathbf{N}^k$.

[2] These functions are usually called Turing computable functions, with a different definition being given for recursive functions. The equivalence of the two definitions is a significant result, but the proof is tedious. The reader is referred to §1 of Chapter X for further information, and to [10], pp. 120–121, 207–237 for full details.

Exercises

3.2. $f: U \to \mathbf{N}^\ell$ is a partial recursive function with domain $U \subseteq \mathbf{N}^k$. Show that there is a Turing machine M such that $\Psi_M^{(k,\ell)} = f$ and such that, for each $(n_1, \ldots, n_k) \in U$, the computation $d_0, d_1, \ldots, d_t$ by M which begins with $d_0 = q_0 \operatorname{code}(n_1, \ldots, n_k)$ ends with $d_t = q \operatorname{code} f(n_1, \ldots, n_k)$ for some internal state q of M.

3.3. Prove that the composition of (partial) recursive functions is (partial) recursive.

3.4. The real number r has decimal expansion $t = r_0 \cdot r_1 r_2 r_3 \cdots$. Given that the function $f: \mathbf{N} \to \mathbf{N}$ defined by $f(n) = r_n$ is not recursive, prove that r is transcendental.

3.5. A subset U of $\mathbf{N}$ is called *recursively enumerable* if it is the range of a recursive function $f: \mathbf{N} \to \mathbf{N}$, or else is empty. Show that $U \subseteq \mathbf{N}$ is recursively enumerable if and only if it is the domain of a partial recursive function.

3.6. A subset U of $\mathbf{N}$ is called *recursive* if its characteristic function is recursive. Prove that $U \subseteq \mathbf{N}$ is recursive if and only if both U and $\mathbf{N} - U$ are recursively enumerable.

3.7. Write a FORTRAN program for calculating the greatest common divisor of two integers of unlimited size (possibly beyond the storage capacity of the machine), assuming the availability of unlimited magnetic tape.

§4 Gödel Numbers

We are interested in delimiting the scope of computations performable by Turing machines, and we are also interested in using Turing machines to formalise the notion of decidability for a logical or mathematical system. To do these things, we need some way of listing all the essentially different Turing machines. From the definition of a Turing machine, it is clear that if two machines M, M' differ only in the labels given their internal states and their alphabets (i.e., if there are bijective maps $\mathfrak{Q} \to \mathfrak{Q}'$, $\mathfrak{S} \to \mathfrak{S}'$ which extend naturally to a bijection $M \to M'$), then M and M' perform essentially the same computations (i.e., the bijection $M \to M'$ extends to a bijection between the sets of computations of M and M'). We may therefore suppose that all Turing machines have alphabets chosen from the universal alphabet $\mathfrak{S}^* = \{s_i \mid i \in \mathbf{N}\}$ (with s_0 corresponding to "blank"), and also that they have lists of internal states chosen from the universal list $\mathfrak{Q}^* = \{q_i \mid i \in \mathbf{N}\}$ (with q_0 corresponding to "initial internal state"). Each machine uses a finite subset of $\mathfrak{Q}^*$, containing q_0, and a finite subset of $\mathfrak{S}^*$, containing s_0. Hence we may think of a Turing machine M as a finite subset of $\mathfrak{Q}^* \times \mathfrak{S}^* \times (\mathfrak{S}^* \cup \{L, R\}) \times \mathfrak{Q}^*$. Further, any tape written in the alphabet $\mathfrak{S}^*$ may be used on an arbitrary Turing machine, for a machine will stop if it scans some symbol not in its alphabet.

We now attach to each Turing machine M a number, called the *Gödel* *number* of M. Denote an element $(a, b, c, d) \in M$ by the string *abcd*. The strings of M have a natural lexicographic order, and by taking all the strings of M in this order we associate with M a unique finite sequence of strings of symbols. We shall define the Gödel number $G(M)$ of M by defining in turn Gödel numbers for every symbol, for every string of symbols and lastly for every finite sequence of strings of symbols.

Define a function $G:\{L, R\} \cup \mathfrak{S}^* \cup \mathfrak{Q}^* \to \mathrm{N}$ by

$$G(L) = 1, G(R) = 3, G(s_i) = 4i + 5, G(q_j) = 4j + 7 (i, j \in \mathrm{N}).$$

If now the symbols a_i have Gödel numbers $G(a_i) = n_i$ $(i = 1, \ldots, r)$, then we define the Gödel number of the string $a_1 \cdots a_r$ by

$$G(a_1 \cdots a_r) = p_1^{n_1} \cdots p_r^{n_r},$$

where p_k denotes the kth prime (so that $p_1 = 2$, $p_2 = 3, \ldots$). The empty string has no Gödel number attached to it.

If $\sigma_1, \ldots, \sigma_s$ are strings of symbols, then we define the Gödel number of the sequence $\sigma_1, \ldots, \sigma_s$ by

$$G(\sigma_1, \ldots, \sigma_s) = p_1^{G(\sigma_1)} p_2^{G(\sigma_2)} \cdots p_s^{G(\sigma_s)}.$$

Finally, the Gödel number of the Turing machine M is defined to be the Gödel number of the unique finite sequence of strings associated with M in the way described before.

Exercises

(In many subsequent exercises, the reader is required to construct a Turing machine. The reader is asked to interpret this as follows: he should convince himself that the required machine can be constructed (perhaps by using previously constructed machines or the results of previous exercises), rather than formally construct the machine as a set of quadruples.)

4.1. Show that, provided each symbol a_j is distinguished from the one element string a_j, and each string σ is distinguished from the sequence σ of length one, then G as defined above is an injective function whose range is a proper subset of N.

4.2. Given a non-empty string σ not containing the symbol s_0, construct a Turing machine which computes $G(\sigma)$ from the initial state $d_0 = q_0\sigma$.

4.3. $f: \mathrm{N} \to \mathrm{N}$ is defined by

$f(n) = 0$ if n is not a Gödel number,
$f(n) = 1$ if n is the Gödel number of a symbol,
$f(n) = 2$ if n is the Gödel number of a string,
$f(n) = 3$ if n is the Gödel number of a finite sequence of strings.

Show that f is recursive.

4.4. The function $f: \mathrm{N} \to \mathrm{N}$ is defined by

$f(n) = 0$ if n is not the Gödel number of a Turing machine,
$f(n) = 1$ if n is the Gödel number of a Turing machine.

Show that f is recursive.

4.5. Turing machines can be ordered by the size of their Gödel numbers. Let $f(n)$ be the Gödel number of the $(n + 1)$-th Turing machine. Show that $f: \mathrm{N} \to \mathrm{N}$ is recursive.

4.6. Use cardinality considerations to prove that there exists a non-recursive function $f: \mathrm{N} \to \mathrm{N}$.

4.7. Show that there is a Turing machine U with the property that, for each Turing machine M and shortest description d of an initial state, U started in the state q code$(G(M), G(d))$

(i) fails if M fails in the state d,

(ii) computes $G(d_t)$ if d_t is the shortest description of the terminal state reached by M starting from d.

A machine such as U is called a universal Turing machine.

In order to apply Turing machines to questions of decidability of mathematical theories, we must be able to encode elements of the appropriate algebras of propositions. We do this by again constructing a universal alphabet and then defining more Gödel numbers. As we can only hope to code countable theories, we confine our attention to them. For each $i \in \mathrm{N}$, the subset $\mathscr{R}_i$ of the set $\mathscr{R}$ of relations of any countable theory is at most countable, so we take a universal set $\mathscr{R}^* = \{r_{ij} | i, j \in \mathrm{N}\}$ of relation symbols, where, for each j, $r_{ij} \in \mathscr{R}_i^*$. Likewise, we take a set $C^* = \{c_j | j \in \mathrm{N}\}$ of constants, and a set $X^* = \{x_j | j \in \mathrm{N}\}$ of variables, and put $V^* = C^* \cup X^*$. For operation symbols we take F, $\Rightarrow$ and $\{(\forall x_j) | j \in \mathrm{N}\}$. We now have a universal alphabet in which every countable theory can be written. Each element of the algebra $P(V^*, \mathscr{R}^*)$ of such a theory has a representative which can be written as a finite string of symbols of this alphabet, for we can replace any $(\forall c_j)$ which occurs, and brackets are unnecessary—we write $a \Rightarrow b$ as $\Rightarrow ab$, $r_{2j}(x_1, x_2)$ as $r_{2j}x_1x_2$, etc. Each string of symbols then has at most one meaning as an element of $\tilde{P}(V^*, \mathscr{R}^*)$.

Exercise 4.8. Prove that each string of symbols has at most one meaning as an element of $\tilde{P}(V^*, \mathscr{R}^*)$.

Gödel numbers are now assigned to our universal alphabet as follows:

$$G(F) = 2, G(\Rightarrow) = 3, G(r_{ij}) = 5^{i+1}7^{j+1}, G(c_j) = 11^{j+1},$$
$$G(x_j) = 13^{j+1}, G((\forall x_j)) = 17^{j+1}.$$

For a string $a_1a_2 \cdots a_n$ of symbols, we put $G(a_1a_2 \cdots a_n) = p_1^{G(a_1)}p_2^{G(a_2)} \cdots p_n^{G(a_n)}$, with p_i denoting the ith prime, as before. For sequences of strings, we also use the method given before. Finally, we define the Gödel number of an element $p \in P(V^*, \mathscr{R}^*)$ to be the least number which is the Gödel number of an element $w \in \tilde{P}(V^*, \mathscr{R}^*)$ which represents p.

Exercises

4.9. Our definitions of Gödel numbers make it possible for an integer to be a Gödel number in the Turing machine sense and also in the propositional algebra sense—e.g., $11 = G(q_1) = G(c_0)$. Modify our definitions of Gödel numbers so that each $n \in \mathrm{N}$ is a Gödel number in at most one way.

4.10. Show that the function $f\colon \mathrm{N} \times \mathrm{N} \to \mathrm{N}$ defined by $f(m, n) = 1$ if m, n are Gödel numbers of elements of $\tilde{P}(V^*, \mathscr{R}^*)$ which represent the same element of $P(V^*, \mathscr{R}^*)$, $f(m, n) = 0$ if m, n are Gödel numbers of elements of $\tilde{P}(V^*, \mathscr{R}^*)$ which represent different elements of $P(V^*, \mathscr{R}^*)$, $f(m, n) = 2$ if either of m, n is not the Gödel number of an element of $\tilde{P}(V^*, \mathscr{R}^*)$, is a recursive function.

§5 Insoluble Problems in Mathematics

We consider various ways in which a mathematical problem can be insoluble, and we begin with two well-known examples—the classical problem of trisecting an angle, and the problem of solving quintic equations. The trisection problem is insoluble by Euclidean construction, but admits a simple solution if a quite minor extension of method is permitted (see [14]). Although there is no formula for the solution of quintic equations by radicals, there is one in terms of elliptic functions (see [5]). Clearly, insolubility of a particular problem depends on a precise statement as to what constitutes a solution.

Each of the above problems is in fact a family of problems. Since a right angle can be trisected, not every angle is impossible to trisect. There exist quintic equations whose solutions are expressible in terms of radicals. The trisection problem asks for a construction which works for every angle, and the non-existence of such a construction follows from the proof that an angle of $\pi/3$ cannot be trisected. Likewise, the existence of a single quintic equation that is insoluble by radicals suffices to demonstrate the non-existence of a general solution by radicals of quintic equations.

Our concern is with the problem of determining for a mathematical theory $\mathscr{T}$ whether or not elements $p_n \in \mathscr{L}(\mathscr{T})$ are theorems of $\mathscr{T}$. In the case of a single element $p \in \mathscr{L}(\mathscr{T})$, let us consider what would constitute a solution to our problem. If p actually is a theorem of $\mathscr{T}$, then we must show that there is a proof of p within $\mathscr{T}$. A proof of p would clearly suffice, provided we can check that it really is a proof. An alleged proof involves only finitely many symbols of our universal alphabet. We can test if a particular step is obtained from earlier steps by modus ponens, or if it is a logical axiom. (We can devise a Turing machine for the purpose.) We could also test the use of Generalisation if we could identify the mathematical axioms of $\mathscr{T}$. In short, proof checking can be performed by a Turing machine provided it can test for mathematical axioms.

Definition 5.1. Let $\mathscr{T}$ be a countable theory expressed in the universal alphabet. We say that $\mathscr{T}$ is *effectively axiomatised* if the characteristic function of the set of Gödel numbers of mathematical axioms of $\mathscr{T}$ is recursive.

This means that if $\mathscr{T}$ is effectively axiomatised, then there is a Turing machine which, when given the Gödel number of an element $q \in P(V^*, \mathscr{R}^*)$, tells us whether or not q is an axiom of $\mathscr{T}$. From the discussion above, it follows that for an effectively axiomatised theory $\mathscr{T} = (\mathscr{R}, A, C)$, there is a Turing machine which, when given the Gödel numbers of $p \in \mathscr{L}(\mathscr{T})$ and of the sequence $p_1, p_2, \ldots, p_n$ of elements of $P(V, \mathscr{R})$, tells us if $p_1, p_2, \ldots, p_n$ is a proof of p in $\mathscr{T}$. Furthermore, there is a Turing machine which, when given the Gödel number of a theorem p of $\mathscr{T}$, computes the smallest number which is the Gödel number of a proof of p.

Suppose now that p is not a theorem of $\mathscr{T}$. If it is the case that $\sim p$ is a theorem of $\mathscr{T}$, then finding a proof of $\sim p$ will not by itself solve our problem, because we would also have to show that $\mathscr{T}$ is consistent, i.e., that F is not a theorem of $\mathscr{T}$. However, if p is not a theorem of $\mathscr{T}$, then the theory $\mathscr{T}' = (\mathscr{R}, A \cup \{\sim p\}, C)$ is consistent, and hence has a model. The construction of a model of $\mathscr{T}'$ would clearly show that p is not a theorem of $\mathscr{T}$. Thus, for any effectively axiomatised theory $\mathscr{T}$ and any p, the problem of deciding whether or not p is a theorem $\mathscr{T}$ is soluble: find a proof of p if p is a theorem, or a model of $\mathscr{T}$ in which p is false if p is not a theorem. Of course, we have not given a general procedure for finding the solution—that is a different problem.

We now consider the case of a family $\{p_n | p \in \mathrm{N}\}$ of propositions of $\mathscr{T}$. The minimal requirement of a solution to the decision problem for the family is clearly that we should know for each n whether or not p_n is a theorem of $\mathscr{T}$. This requirement can be met by simply requiring the solution to be the determination of the function $f\colon \mathrm{N} \to \{0, 1\}$ such that $f(p_n) = 1$ if and only if p_n is a theorem. For the determination to be satisfactory, the function f must be capable of calculation in some routine manner, which means that f must be a recursive function. For this to be so, the family $\{p_n | n \in \mathrm{N}\}$ must be able to be systematically computed, i.e., there is a condition on the family in order that our decision problem be well posed. With these considerations in mind, we make the following definitions.

Definition 5.2. The family $\{p_n | n \in \mathrm{N}\}$ is called *recursively enumerable* if $\{G(p_n) | n \in \mathrm{N}\}$ is a recursively enumerable subset of N, and it is *recursive* if $\{G(p_n) | n \in \mathrm{N}\}$ is a recursive subset of N. If $G(p_n)$ is a recursive function of n, the family is called *recursively enumerated.*

Definition 5.3. Let $\mathscr{F} = \{p_n | n \in \mathrm{N}\}$ be a recursively enumerated family of propositions of the theory $\mathscr{T}$. We say the decision problem for $\mathscr{F}$ is *recursively soluble* if the characteristic function of $\{n \in \mathrm{N} | \mathscr{T} \vdash p_n\}$ is recursive.

If $\mathscr{T}$ is a countable theory (written in the universal alphabet), then the Gödel numbering of elements of $\mathscr{L}(\mathscr{T})$ orders $\mathscr{L}(\mathscr{T})$ and so provides a recursive enumeration of $\mathscr{L}(\mathscr{T})$. The theory $\mathscr{T}$ is then called *decidable* if the family $\mathscr{L}(\mathscr{T})$ has recursively soluble decision problem.

Our decidability criterion is based on the minimum answer we could expect—a "yes/no" answer. We do not require the Turing machine which provides this answer also to prove it by giving either a proof or a countermodel for each p_n. If the present formulation of a solution produces undecidable theories, then any more rigorous requirement must be expected to render even more problems insoluble.

Exercise 5.4. The consistent theory $\mathscr{T}$ is effectively axiomatised and complete. Show that $\mathscr{T}$ is decidable. Show further that there is a Turing machine which, given the Gödel number of an element $p \in \mathscr{L}(\mathscr{T})$, answers the question of whether or not p is a theorem, and also provides a proof of its answer.

The notion that a family of objects can have a recursively insoluble decision problem of some kind can be applied to situations in our informal mathematics, as the following example shows. Later, we shall find examples within formal mathematical structures.

Example 5.5. Let M_n denote the nth Turing machine. The problem is to determine for all integers n, r, whether or not there is a computation of M_n beginning with the state q_0 code(r). I.e., the problem is to determine whether or not an arbitrary Turing machine M_n, fed with an arbitrary integer r, stops. We show that this stopping problem is recursively insoluble. Put $f_n = \Psi_{M_n}^{(1,1)}$. The problem is to determine those (n, r) for which $f_n(r)$ exists, and we show that there is no Turing machine which computes for each n whether or not $f_n(n)$ exists. Suppose the function $f\colon \mathrm{N} \to \mathrm{N}$ defined by

$$\begin{aligned} f(n) &= 1 \quad \text{if} \quad f_n(n) \text{ exists}, \\ f(n) &= 0 \text{ otherwise}, \end{aligned}$$

is recursive. Then the function $g\colon \mathrm{N} \to \mathrm{N}$ defined by

$$\begin{aligned} g(n) &= f_n(n) + 1 \quad \text{if} \quad f_n(n) \text{ exists}, \\ g(n) &= 0 \text{ otherwise}, \end{aligned}$$

is also recursive, since $f_n(n)$ can be computed when it exists. We now have a contradiction, for since $\{f_n | n \in \mathrm{N}\}$ contains the set of recursive functions, $g = f_m$ for some integer m, and then

$$f_m(m) = g(m) = f_m(m) + 1.$$

Hence f is not recursive, and so there is no Turing machine which determines for all n, whether or not $f_n(n)$ exists. In fact, since $h(n) = f_n(n)$ is partial recursive, there is a Turing machine M which computes it, and we have proved that M has a recursively insoluble stopping problem.

Exercises

5.6. Use $h(n) = f_n(n)$ to construct a recursively enumerable set E which is not recursive.

5.7. Let $A = \{n \in \mathbf{N} | f_n \text{ is recursive}\}$. Prove that A is not recursively enumerable.

5.8. Let E be a recursively enumerable subset of N. Show that there exists a Turing machine which, started in the state $q_0 s_1^n$, stops with blank tape if $n \in E$, and fails to stop if $n \notin E$. (Hint: if f: N → N is a recursive function with $f(\mathbf{N}) = E$, then, for given n, compute in turn $f(0), f(1), \ldots$ until the first r (if any) for which $f(r) = n$ is found.)

§6 Insoluble Problems in Arithmetic

In §5 we gave an example of a family of objects in informal mathematics for which a decision problem is recursively insoluble. We now wish to convert this example into an example within formal arithmetic. We do this in a way which will allow us to apply some of our ideas and results to other interesting systems. For that reason, we shall be concerned with theories which formalise some aspects of the theory of N (which is our underlying object of study) and it is convenient to set down first some notational conventions and some definitions. Throughout this section, θ, s, a, m respectively denote the property of being 0, the successor relation, the addition relation and the multiplication relation. Whenever N is given as a model of a theory $\mathscr{T}$, it is understood that any of θ, s, a, m which are relation symbols of $\mathscr{T}$ have their standard interpretations. Axioms which we will use in our constructions are

1) the Peano axioms P_1, P_2, P_3, P_4, P_5 of the first-order theory $\mathscr{P}$ of §4 of Chapter VI. (Recall that the scope of the axiom scheme of induction (P_5) depends on the theory under consideration.)

2) the addition axioms

$$\begin{aligned}
\text{add}_1 &= (\forall x)(\forall y)(\exists! z)a(x, y, z),\\
\text{add}_2 &= (\forall x)(\forall y)(\theta(y) \Rightarrow a(x, y, x)),\\
\text{add}_3 &= (\forall x)(\forall y)(\forall z)(\forall t)(\forall u)(s(z, y) \wedge a(x, z, t)\\
&\qquad \wedge a(x, y, u) \Rightarrow s(t, u)).
\end{aligned}$$

3) the multiplication axioms

$$\begin{aligned}
\text{mult}_1 &= (\forall x)(\forall y)(\exists! z)m(x, y, z),\\
\text{mult}_2 &= (\forall x)(\forall y)(\theta(y) \Rightarrow m(x, y, y))\\
\text{mult}_3 &= (\forall x)(\forall y)(\forall z)(\forall t)(\forall u)(s(z, y) \wedge m(x, z, t)\\
&\qquad \wedge m(x, y, u) \Rightarrow a(u, x, t)).
\end{aligned}$$

4) for theories with N contained in the set of constants, the identification axioms

$$\begin{aligned}
e_n = (\exists x_0)(\exists x_1) \cdots (\exists x_{n-1})(\theta(x_0) \wedge s(x_1, x_0) \wedge \cdots \wedge\\
s(x_{n-1}, x_{n-2}) \wedge s(n, x_{n-1})).
\end{aligned}$$

We shall deal mainly with the theory $\mathscr{N}$ with relation symbols $\mathscr{R}(\mathscr{N}) = \{=, \theta, s, a, m\}$, constants N, and axioms $A(\mathscr{N})$ being all those listed above.

We call $\mathscr{N}$ recursive arithmetic. To assist us, we shall also use the theory $\mathscr{N}_0$, which differs from $\mathscr{N}$ only in that the axiom scheme of induction is excluded from the axioms. Both $\mathscr{N}$ and $\mathscr{N}_0$ are effectively axiomatised. As we need to compare theories, we make the following definitions.

Definition 6.1. Let $\mathscr{T} = (\mathscr{R}, A, C)$ and $\mathscr{T}' = (\mathscr{R}', A', C')$ be first-order theories. We say $\mathscr{T}'$ *extends* $\mathscr{T}$, and write $\mathscr{T}' \supseteq \mathscr{T}$, if $\mathscr{R}' \supseteq \mathscr{R}$, $A' \supseteq A$ and $C' \supseteq C$.

Definition 6.2. Let $\mathscr{T}' \supseteq \mathscr{T}$, and let $M = (M, v, \psi)$ be a model of $\mathscr{T}$. We say that M *extends to a model of* $\mathscr{T}'$ if there exist $v' : C' \to M$ and $\psi' : \mathscr{R}' \to \mathrm{rel}(M)$, extending v, ψ respectively, such that (M, v', ψ') is a model of $\mathscr{T}'$.

Definition 6.3. Let $\mathscr{T} = (\mathscr{R}, A, C) \supseteq \mathscr{N}_0$ and have N as model. Let $f : U \to \mathrm{N}$ be a function defined on some subset U of N. We say that f is *strongly definable* in $\mathscr{T}$ if there is an element $p(x, y) \in P(V, \mathscr{R})$ such that, for all $m, n \in \mathrm{N}$, $\mathscr{T} \vdash p(m, n)$ if and only if $m \in U$ and $f(m) = n$. The definition is extended in the obvious way for functions of several variables.

The key result we intend to prove is that if $\mathscr{T} \supseteq \mathscr{N}_0$ and has N as model, then every partial recursive function is strongly definable in $\mathscr{T}$. The proof is tedious, although the idea is simple—we build up descriptions of the state of a given Turing machine as a function of the input and the number of steps performed.

Definition 6.4. The *state function* corresponding to the state $[s_{\beta_\ell} \cdots s_{\beta_1} q_i s_{\alpha_1} \cdots s_{\alpha_k}]$ is the function $f : \mathrm{N} \to \mathrm{N}$ given by

$$\begin{aligned} f(0) &= G(q_i), \\ f(2i + 1) &= G(s_{\alpha_{i+1}}),\ 0 \leqslant i \leqslant k - 1, \\ f(2i + 1) &= G(s_0),\ i \geqslant k, \\ f(2i + 2) &= G(s_{\beta_{i+1}}),\ 0 \leqslant i \leqslant \ell - 1, \\ f(2i + 2) &= G(s_0),\ i \geqslant \ell. \end{aligned}$$

State functions are always strongly definable, as they take the value $G(s_0)$ except on a finite set. For a given Turing machine, it is easy to construct a description of the state function f_1 produced from an initial state f after one step of the computation. Continuing, one can produce, for any n, a description of the state function f_n after n steps. The difficulty in this approach is that the complexity of the description so obtained increases with n, whereas we need a single description of $f_x(y)$ as a function of the two variables x, y. Fortunately, there is a trick which allows us to give bounded definitions of arbitrary finite sequences.

Lemma 6.5. (The Sequence Number Lemma). *There exists a strongly definable function* $\mathrm{seq} : \mathrm{N}^+ \times \mathrm{N} \to \mathrm{N}$ *such that, for any n and $a_0, a_1, \ldots, a_n \in \mathrm{N}$, there exists $b \in \mathrm{N}^+$ with the property that* $\mathrm{seq}(b, r) = a_r$ *for* $r = 0, \ldots, n$.

Proof. Let $T(n)$ denote the nth triangular number:

$$T(n) = 1 + 2 + \cdots + n = \tfrac{1}{2}n(n + 1).$$

For each $z > 0$, there is a unique n such that

$$T(n) < z \leqslant T(n + 1) = T(n) + n + 1.$$

Thus z is uniquely expressible as $z = T(n) + y$ with $0 < y \leqslant n + 1$. (We choose this range for y because later we shall need $y \neq 0$.) Put $x = n + 2 - y$. Then x, y are uniquely determined functions of z, which we denote by $L(z)$, $R(z)$ respectively. Put $P(x, y) = T(x + y - 2)$. P, L, R are strongly definable functions, for we may regard $z = P(x, y)$ as an abbreviation for

$$(x > 0) \wedge (y > 0) \wedge (2z = (x + y - 2)(x + y - 1) + 2y),$$

$x = L(z)$ as one for

$$(x > 0) \wedge (z > 0) \wedge (\exists y)((y > 0) \wedge (2z = (x + y - 2)(x + y - 1) + 2y)),$$

and $y = R(z)$ as one for

$$(y > 0) \wedge (z > 0) \wedge (\exists x)((x > 0) \wedge (2z = (x + y - 2)(x + y - 1) + 2y)).$$

The function $\operatorname{seq}(b, r)$ is defined to be the remainder on division of $L(b)$ by $1 + (r + 1)R(b)$. This is strongly definable, the relation $z = \operatorname{seq}(x, y)$ being given by

$$(x > 0) \wedge (z < 1 + (y + 1)R(x)) \wedge (\exists t)(L(x) = t(1 + (y + 1)R(x)) + z).$$

Finally, given $a_0, a_1, \ldots, a_n \in \mathrm{N}$, we have to find $b \in \mathrm{N}^+$ such that $\operatorname{seq}(b, r) = a_r$ for $0 \leqslant r \leqslant n$. Pick $c \in \mathrm{N}$ such that $c > a_r$ for $0 \leqslant r \leqslant n$ and such that c is divisible by each of $1, 2, \ldots, n$. Put $m_r = 1 + (r + 1)c, r = 0, \ldots, n$. m_r and m_s are relatively prime for every pair r, s such that $0 \leqslant r < s \leqslant n$, for if d is a common divisor of m_r and m_s, d also divides $(s + 1)m_r - (r + 1)m_s = s - r$. Hence d divides c, and the definition of m_r shows now that $d = 1$. We may therefore apply the Chinese Remainder Theorem (see [10], p 135) to the system of congruences

$$x \equiv a_r \bmod m_r\ (r = 0, \ldots, n).$$

Let e be a positive solution to this system, and put $b = P(e, c)$. Then $e = L(b)$, $c = R(b)$, $L(b) \equiv a_r \bmod(1 + (r + 1)R(b))$, and $a_r < c < 1 + (r + 1)R(b)$, showing that $a_r = \operatorname{seq}(b, r)$ for $r = 0, \ldots, n$. □

Exercises

6.6. Given $m, n, r \in \mathrm{N}$ such that $m + n = r$, prove that $\mathcal{N}_0 \vdash a(m, n, r)$. Hence show that if $\mathcal{T} \supseteq \mathcal{N}_0$ and has N as a model, then $\mathcal{T} \vdash a(m, n, r)$ implies $\mathcal{N}_0 \vdash a(m, n, r)$ for $m, n, r \in \mathrm{N}$. Do the same thing for multiplication.

6.7. For m, n, $r \in \mathrm{N}$ and $\mathcal{T} \supseteq \mathcal{N}_0$ with N as model, show that $\mathcal{T} \vdash \operatorname{seq}(m, n) = r$ if and only if $\operatorname{seq}(m, n) = r$. (This shows that the formula given above as a definition in $\mathcal{T}$ of seq indeed strongly defines seq.)

The sequence number function defined in Lemma 6.5 enables us to give definitions in $\mathcal{T}$ of various functions describing a computation by a Turing machine M. We give the definitions and leave the reader to verify them. If M has a quadruple $(q_\alpha, s_\beta, a, b)$, we define $M_{\alpha,\beta}(x, y, z) \in P(V, \mathcal{R})$ as follows. We have $b = q_\gamma$ for some γ, $a = s_{\beta'}$ (for some β') or $a = L$ or $a = R$. Put

$$M_{\alpha,\beta}(x, y, z) = (\operatorname{seq}(x, 0) = G(q_\alpha)) \wedge (\operatorname{seq}(x, 1) = G(s_\beta)) \\ \wedge (y = 0 \Rightarrow z = G(q_\gamma)) \wedge K(x, y, z),$$

where

$$K(x, y, z) = (y = 1 \Rightarrow z = G(s_{\beta'})) \wedge (y > 1 \Rightarrow z = \operatorname{seq}(x, y)) \quad \text{if} \quad a = s_{\beta'},$$
$$K(x, y, z) = [((\exists k)(y = 2k + 1)) \Rightarrow z = \operatorname{seq}(x, y + 2)] \wedge (y = 2 \Rightarrow z = \operatorname{seq}(x, 1)) \\ \wedge [((\exists k)(y = 2k + 4)) \Rightarrow z = \operatorname{seq}(x, y - 2)] \quad \text{if} \quad a = R,$$
$$K(x, y, z) = (y = 1 \Rightarrow z = \operatorname{seq}(x, 2) \wedge [((\exists k)(y = 2k + 3)) \Rightarrow z = \operatorname{seq}(x, y - 2)] \\ \wedge [((\exists k)(y = 2k + 2)) \Rightarrow z = \operatorname{seq}(x, i + 2)] \quad \text{if} \quad a = L.$$

Now put

$$M(x, y, z) = \bigvee_{\alpha,\beta} M_{\alpha,\beta}(x, y, z),$$

where the disjunction is taken over the finitely many pairs α, β for which there is a quadruple $(q_\alpha, s_\beta, a, b) \in M$. If there are no such quadruples, put $M(x, y, z) = F$.

Suppose that f, g are state functions such that $[f] \overset{M}{\to} [g]$. For $r \in \mathrm{N}$, let $u \in \mathrm{N}$ be such that $\operatorname{seq}(u, i) = f(i)$ for $i = 0, \ldots, r + 2$. If $k \in \mathrm{N}$, we claim that $k = g(r)$ if and only if $\mathcal{T} \vdash M(u, r, k)$. We can now prove some results.

Lemma 6.8. *Let f be an initial state function (i.e., $f(0) = G(q_0)$) and let $g(n, r)$ be the value at r of the state function after n steps of the computation by M starting at $[f]$. Then g is strongly definable in $\mathcal{T}$.*

Proof. f is strongly definable, and so we give a definition of g in terms of the definition of f. Put

$$\varphi(x, y, z) = (\exists u)[(\forall v)(v \leqslant y + 2x \Rightarrow \operatorname{seq}(\operatorname{seq}(u, 0), v) = f(v)) \\ \wedge (\operatorname{seq}(\operatorname{seq}(u, x), y) = z) \wedge (\forall w)(\forall t)(((1 \leqslant w \leqslant x) \\ \wedge (t \leqslant y + 2(x - w))) \Rightarrow M(\operatorname{seq}(u, w - 1), t, \operatorname{seq}(\operatorname{seq}(u, w), t)))]$$

Then $g(n, r) = k$ if and only if $\mathcal{T} \vdash \varphi(n, r, k)$, whence the result. □

Any initial state function f can be expressed in terms of two integers u, v, since we can always find u, v such that $f(x) = \operatorname{seq}(u, x)$ if $x \leqslant v$, and $f(x) = G(s_0)$ if $x > v$. (If so desired, we can replace u, v by the single integer $w = P(u, v)$, using $u = L(w)$, $v = R(w)$.) If this definition of f is substituted into the element φ given above, we obtain a 5-variable formula, $\psi(u, v; x, y, z)$ say, which describes the behavior of M for any input. We can express the statement that M, started in the state given by (u, v), stops in the state of the function whose value at y is z, by

$$(\exists x)(\psi(u, v; x, y, z) \wedge (\forall x')(x' > x \Rightarrow (\forall t)(\sim\psi(u, v; x', y, t)))).$$

Theorem 6.9. *Let $\mathscr{T} \supseteq \mathscr{N}_0$ be a theory with* N *as model. Then every partial recursive function is strongly definable in $\mathscr{T}$.*

Proof. The formulae given above, together with a description of the input function in terms of the arguments of $\Psi_M^{(k,\ell)}$, can be adapted to give a definition of $\Psi_M^{(k,\ell)}$. The reader is asked to supply the details. □

We are now able to provide an example of an insoluble decision problem within the formal theory $\mathscr{N}$. From Theorem 6.9, it follows that any relation on N whose characteristic function is recursive is also strongly definable in any theory $\mathscr{T}$ of the type considered above. In particular, there is a formula, $\text{comp}(x_1, x_2, x_3)$ say, defining the relation that the machine of Gödel number x_1, applied to the number x_2, computes x_3. Reference to Example 5.5 shows that the family $\{(\exists x)\text{comp}(n, n, x) | n \in \text{N}\}$ has an insoluble decision problem.

Theorem 6.10. *Let $\mathscr{T} \supseteq \mathscr{N}_0$ be a theory which has* N *as model. Then $\mathscr{T}$ is undecidable. In particular, $\mathscr{N}$ is undecidable.*

Proof. A decision process for $\mathscr{T}$ would provide a decision process for the family $\{(\exists x)\text{comp}(n, n, x) | n \in \text{N}\}$. □

Theorem 6.11. *Let $\mathscr{T} \supseteq \mathscr{N}_0$ be an effectively axiomatised theory with* N *as model. Then $\mathscr{T}$ is incomplete.*

Proof. By Exercise 5.4 and Theorem 6.10. However, it is of interest to construct an element $q \in \mathscr{L}(\mathscr{T})$ such that neither q nor $\sim q$ is a theorem of $\mathscr{T}$. Let $\mathscr{T} = (\mathscr{R}, A, C)$, and write P for $P(V, \mathscr{R})$. Let $G: P \to \text{N}$ denote the Gödel number function, and let $F: G(P) \to P$ denote its inverse (G is injective). Since proofs in $\mathscr{T}$ can be checked by Turing machine, the relation "x_2 is the Gödel number of a proof in $\mathscr{T}$ of $F(x_1)$" is recursive. Let $\text{proof}_{\mathscr{T}}(x_1, x_2)$ be a definition of this relation in $\mathscr{T}$, and put

$$\text{theorem}_{\mathscr{T}}(x_1) = (\exists x_2)\text{proof}_{\mathscr{T}}(x_1, x_2).$$

Then theorem $_{\mathscr{T}}(x_1)$ defines in $\mathscr{T}$ the property "x_1 is the Gödel number of a theorem of $\mathscr{T}$".

For any element $w \in P$, write $w(x_0)$ to denote its (possible) dependence on x_0. If $n \in \text{N}$, then $n \in C$ and so $w(n) \in P$. We consider $w(n)$ as a function of both n and w, and denote it by $\text{sub}(n, w)$. Define $\varphi(m, n) = G(\text{sub}(m, F(n)))$, for $m \in \text{N}$ and $n \in G(P)$. φ is then a partial recursive function, hence there is an element $p(x_1, x_2, x_3) \in P$ defining the relation $\varphi(x_1, x_2) = x_3$.

We now put

$$\pi(x_1, x_2) = (\exists x_3)(p(x_1, x_2, x_3) \wedge \text{theorem}_{\mathscr{T}}(x_3)),$$

and consider the meaning of $\pi(x_1, x_2)$ in certain cases. If w satisfies $\text{var}(w) \subseteq \{x_0\} \cup C$, then $w(m) \in \mathscr{L}(\mathscr{T})$ for all $m \in \text{N}$. Choose such a w, and let $n = G(w)$. Then $\pi(m, n)$ is true in N if and only if, for some $a \in \text{N}$, we have both $\varphi(m, n) = a$

and a is the Gödel number of a theorem of $\mathscr{T}$. Since $\varphi(m, n) = G(w(m))$, we see that a must be both the Gödel number of $w(m)$ and the Gödel number of a theorem. Hence $\pi(m, n)$ is true in N if and only if $\mathscr{T} \vdash w(m)$. We now choose $w(x_0) = \sim\pi(x_0, x_0)$, so that $n = G(\sim\pi(x_0, x_0))$, and put $q = w(n)$. Then $\pi(n, n)$ is true in N if and only if $\mathscr{T} \vdash q$. But $q = \sim\pi(n, n)$, hence $\mathscr{T} \vdash q$ if and only if q is false in N. Since N is a model of $\mathscr{T}$, $\mathscr{T} \vdash q$ implies q is true in N. Hence q cannot be a theorem of $\mathscr{T}$, which from the condition above implies q is true in N, which then implies that $\sim q$ cannot be a theorem of $\mathscr{T}$. Thus $q = \sim\pi(n, n)$ has the property required to demonstrate the incompleteness of $\mathscr{T}$. □

We note that this incompleteness cannot be cured by adding q as an axiom to form a new theory $\mathscr{T}' \supseteq \mathscr{T}$, because replacing $\text{theorem}_{\mathscr{T}}(x_3)$ by $\text{theorem}_{\mathscr{T}'}(x_3)$ in our construction provides another element q' with the requisite properties. The proof shows that no effective axiomatisation of N can lead to a complete theory.

The result of Theorem 6.11 is known as Gödel's Incompleteness Theorem.

Exercises

6.12. Show that $\{n \in \mathrm{N} | n = G(p)$ for some $p \in \mathscr{L}(\mathscr{N})$ true in N$\}$ is not recursively enumerable.

6.13. Show that $\{n \in \mathrm{N} | n = G(p)$ for some $p \in \mathscr{L}(\mathscr{N})$ such that $\mathscr{N} \vdash p\}$ is recursively enumerable but not recursive.

§7 Undecidability of the Predicate Calculus

We investigate the decidability of the predicate calculus by taking a known undecidable theory $\mathscr{T} = (\mathscr{R}, A, C)$, and trying to show that the theory $(\mathscr{R}, \varnothing, \varnothing)$ is also undecidable. The method is to suppose the existence of a decision process for $(\mathscr{R}, \varnothing, \varnothing)$ and to construct from it a decision process for $(\mathscr{R}, A, C)$. The following simple result will be used.

Lemma 7.1. *Let $\mathscr{T}$, $\mathscr{T}'$ be theories, and let $\varphi : \mathscr{L}(\mathscr{T}) \to \mathscr{L}(\mathscr{T}')$ be a recursive function such that for all $p \in \mathscr{L}(\mathscr{T})$, we have $\mathscr{T} \vdash p$ if and only if $\mathscr{T}' \vdash \varphi(p)$. Suppose $\mathscr{T}'$ is decidable. Then $\mathscr{T}$ is decidable.*

Proof: Clearly, to determine if p is a theorem, it suffices to calculate $\varphi(p)$ and to apply the decision process for $\mathscr{T}'$. □

Lemma 7.2. *Let $\mathscr{N}_1$ be the theory formed from the theory $\mathscr{N}_0$ by omitting the constants and the axioms e_n which identify the constants. Then $\mathscr{N}_1$ is undecidable.*

Proof: Put, for each $n \in \mathrm{N}$,

$$e_n(x) = (\exists x_0)(\exists x_1) \cdots (\exists x_{n-1})(\theta(x_0) \wedge s(x_1, x_0) \wedge \cdots \wedge s(x, x_{n-1})).$$

Now for any $p \in \mathscr{L}(\mathscr{N}_0)$, $\text{var}(p) \subseteq \mathrm{N}$. Hence there is an element

$p(x_1, \ldots, x_r) \in P(V - \mathrm{N}, \mathscr{R})$, such that there exist integers $n_1, \ldots, n_r$ for which $p = p(n_1, \ldots, n_r)$. We define $\varphi: \mathscr{L}(\mathscr{N}_0) \to \mathscr{L}(\mathscr{N}_1)$ by

$$\varphi(p) = (\forall x_1) \cdots (\forall x_r)(e_{n_1}(x_1) \wedge \cdots \wedge e_{n_r}(x_r) \Rightarrow p(x_1, \ldots, x_r)).$$

In order to complete the proof by an appeal to the previous lemma, we have to show that $\mathscr{N}_0 \vdash p$ if and only if $\mathscr{N}_1 \vdash \varphi(p)$. Suppose that $\mathscr{N}_1 \vdash \varphi(p)$. Since $\mathscr{N}_0 \supseteq \mathscr{N}_1$, $\mathscr{N}_0 \vdash \varphi(p)$. Since $e_{n_i}(n_i)$ is an axiom of $\mathscr{N}_0$ for all i, it follows immediately that $\mathscr{N}_0 \vdash p(n_1, \ldots, n_r)$, i.e., that $\mathscr{N}_0 \vdash p$. Now suppose $\mathscr{N}_0 \vdash p$. Since $\mathscr{N}_1 \vdash (\exists! x)e_n(x)$ for each $n \in \mathrm{N}$, then for each $n \in \mathrm{N}$ there is in any model M of $\mathscr{N}_1$ a unique element $m_n \in M$ such that $M \models e_n(m_n)$. By mapping n to m_n we make M a model of $\mathscr{N}_0$, so $p(m_{n_1}, \ldots, m_{n_r})$ is true in M. The uniqueness of the m_n now implies that $\varphi(p)$ is true in M. Thus $\varphi(p)$ is true in every model of $\mathscr{N}_1$, and so $\mathscr{N}_1 \vdash \varphi(p)$. □

Lemma 7.3. *Let $V = \{x_0, x_1, \ldots\}$, and $\mathscr{R} = \{\rho\}$, where ρ is a 4-ary relation symbol. Then* Pred$(V, \mathscr{R})$ *is undecidable.*

Proof: Since Pred$(V, \mathscr{R})$ does not involve either the identity relation symbol or the axioms of identity, we first consider these axioms. The theory $\mathscr{N}_1$ has only finitely many relation symbols, hence the axiom scheme of substitution of identical elements is finite. Thus $\mathscr{N}_1$ has only finitely many axioms of identity. Denote the conjunction of all of these axioms by a.

The relation symbols of $\mathscr{N}_1$ will be replaced by ρ, which intuitively is regarded as follows: $\rho(x, y, z, t)$ means $xy + z = t$. Define a homomorphism $f: P(V, \mathscr{R}^{(1)}) \to P(V, \mathscr{R})$, where $\mathscr{R}^{(1)} = \{=, \theta, s, a, m\}$, by

$$\begin{aligned}
f(\theta(x)) &= \rho(x, x, x, x), \\
f(x = y) &= (\forall z)(\forall t)(\rho(z, z, z, z) \Rightarrow \rho(z, t, x, y)), \\
f(s(x, y)) &= (\forall z)(\forall t)((\rho(z, z, z, z) \wedge (\forall u)\rho(t, u, z, u)) \Rightarrow \rho(t, y, t, x)), \\
f(a(x, y, z)) &= (\forall t)((\forall u)(\forall v)(\rho(u, u, u, u) \Rightarrow \rho(t, v, u, v)) \Rightarrow \rho(t, x, y, z)), \\
f(m(x, y, z)) &= (\forall t)(\rho(t, t, t, t) \Rightarrow \rho(x, y, t, z)),
\end{aligned}$$

for all $x, y, z \in V$. Then define $g: P(V, \mathscr{R}^{(1)}) \to P(V, \mathscr{R})$ by $g(p) = f(a) \Rightarrow f(p)$. We show that Lemma 7.1 applies, for then the undecidability of $\mathscr{N}_1$ suffices to complete the proof.

Suppose $\mathscr{N}_1 \vdash p$. A proof of p from a maps under f into a proof of $f(p)$ from $f(a)$. By the Deduction Theorem, $f(a) \Rightarrow f(p)$ is a theorem. Conversely, suppose $f(z) \Rightarrow f(p)$ is a theorem. If M is any model of $\mathscr{N}_1$, then interpreting $\rho(x, y, z, t)$ as $xy + z = t$ gives an interpretation of $P(V, \mathscr{R})$ in which $f(a)$ is true. Since $f(a) \Rightarrow f(p)$ is a theorem, we conclude that $f(p)$ is true. By the way the interpretation of ρ is defined, p is also true in M. Hence p is a theorem of $\mathscr{N}_1$. □

Corollary 7.4. (Church's Theorem) *Let $\mathscr{R}^* = \{r_{ij} | i, j \in \mathrm{N}\}$, with r_{ij} an i-ary relation symbol, be the universal relation alphabet. Then* Pred$(V, \mathscr{R}^*)$ *is undecidable.*

Proof: The inclusion $P(V, \mathscr{R}) \to P(V, \mathscr{R}^*)$ satisfies the conditions of Lemma 7.1. □

We end the chapter with the proof of a stronger result due to Kalmar.

Theorem 7.5. *Let r be a binary predicate symbol. Then* Pred(V, $\{r\}$) *is undecidable.*

Before giving the formal proof, we note that the result implies that if $\mathscr{R}$ contains at least one n-ary relation symbol with $n \geqslant 2$, then Pred(V, $\mathscr{R}$) is undecidable. The theorem will be proved by constructing a function $f: P(V, \{\rho\}) \to P(V, \{r\})$, where ρ is a 4-ary relation symbol, such that $f(p)$ is a theorem if and only if p is a theorem, and in addition such that if $\text{var}(p) = \varnothing$, then $\text{var}(f(p)) = \varnothing$. The construction uses the following idea, which shows how to express a 4-ary relation ρ on a given set S in terms of a binary relation r on a related set S'. (For convenience we shall use ρ, r also to denote interpretations of the relation symbols ρ, r.)

Lemma 7.6. *Let ρ be a 4-ary relation on the non-empty set S. Put $S' = \{K\} \cup S^2 \cup S^4$. For $x \in S$, define $\Delta(x) = (x, x) \in S'$. Let r be the binary relation on S' consisting of those pairs (a, b) for which at least one of the following holds:*

(i) *$a = (x, y), b = (z, t)$, where $x, y, z, t \in S$ and $x = y = z$ or $y = z = t$,*
(ii) *$a = (x, y), b = (x, y, z, t)$ where $x, y, z, t \in S$,*
(iii) *$a = (x, y, z, t), b = (z, t)$ where $x, y, z, t \in S$,*
(iv) *$a = K, b = (x, y, z, t)$ where $(x, y, z, t) \in \rho$.*

Then the elements of $\Delta(S)$, and of ρ, can be characterised in terms of r.

Proof: An element $a \in S'$ is in $\Delta(S)$ if and only if $(a, a) \in r$. We claim that a quadruple (x, y, z, t) of elements of S is in ρ if and only if their images X, Y, Z, T under Δ satisfy the condition that there are elements $A, B, C, D \in S'$ such that all the pairs (X, A), (A, Y), (Z, B), (B, T), (A, C), (C, B) and (D, C) are in r, but (E, D) is not in r for any E. To show this, observe that (E, D) not in r for any E implies that $D = K$. Then (D, C) is in r if and only if C is a quadruple in ρ. (C, B) in r requires B to be the final pair of C. (A, C) in r shows that A is either K or the initial pair of C, and the former is excluded if (X, A) is in r. Hence if the condition is satisfied by X, Y, Z, T, then $C = (x, y, z, t)$ and is in ρ. □

Proof of Theorem 7.5. Put

$$R(x, y, z, t) = (\exists a)(\exists b)(\exists c)(\exists d)(r(x, a) \wedge r(a, y) \wedge r(z, b) \wedge r(b, t) \wedge r(a, c) \wedge r(c, b) \wedge r(d, c) \wedge (\forall e) \sim r(e, d))$$

We define $f: P(V, \{\rho\}) \to P(V, \{r\})$ in terms of a prefix π and a kernel k. If $\text{var}(p) \neq \varnothing$, we define π_p to be the conjunction of the $r(x, x)$ for which

$x \in \mathrm{var}(p)$, while if $\mathrm{var}(p) = \varnothing$, π_p is not defined. The kernel $k(p)$ is defined inductively by

$$\begin{aligned} &k(F) = F, \\ &k(\rho(x, y, z, t)) = R(x, y, z, t) \text{ for all } x, y, z, t \in V, \\ &k(p \Rightarrow q) = k(p) \Rightarrow k(q), \\ &k((\forall x)p) = (\forall x)(r(x, x) \Rightarrow k(p)). \end{aligned}$$

We now put

$$\begin{aligned} &f(p) = \pi_p \Rightarrow k(p) \quad \text{if} \quad \mathrm{var}(p) \neq \varnothing, \\ &f(p) = k(p) \quad \text{if} \quad \mathrm{var}(p) = \varnothing, \end{aligned}$$

and show f satisfies the conditions of Lemma 7.1. Suppose that $\mathrm{var}(p) = \varnothing$ and $f(p)$ is a theorem. The truth or falsity of p in any interpretation depends only on the choice of the set S and of the 4-ary relation ρ on S. We construct S', and r on S', as in Lemma 7.6. Since the definition of f effectively limits consideration to elements of $\Delta(S)$, we find that p is true in S if and only if $f(p)$ is true in S'. Since $f(p)$ is a theorem, we conclude that p is true in every interpretation and so is also a theorem.

Conversely, suppose p is a theorem, and let $p_1, \ldots, p_n$ be a proof of p. We use induction over n to show that $f(p)$ is a theorem. (We do not assume $\mathrm{var}(p) = \varnothing$, as this would upset the induction.) Suppose then that $f(p_1), \ldots, f(p_{n-1})$ are theorems. There are three possibilities for p_n: it is an axiom, it is obtained by modus ponens, or it is obtained by Generalisation.

If p is an axiom, then $f(p)$, although not an axiom, is easily seen to be provable. Similarly, if p follows from p_i and p_j by modus ponens, then (by use of truth functions or the Deduction Theorem) $f(p)$ is deducible from $f(p_i)$ and $f(p_j)$. Suppose finally that $p = (\forall x)q$ is obtained by Generalisation. Then $p_{n-1} = q$, and $f(q) = \pi_q \Rightarrow k(q)$ is a theorem (in the other case, $\mathrm{var}(q) = \varnothing$ and $f(p)$ is trivially a theorem). Since π_q is either π_p or $\pi_p \wedge r(x, x)$, it follows that $\pi_p \Rightarrow (r(x, x) \Rightarrow k(q))$ is a theorem, and Generalisation yields $(\forall x)(\pi_p \Rightarrow (r(x, x) \Rightarrow k(q)))$. Since $x \notin \mathrm{var}(p)$, this implies $\pi_p \Rightarrow (\forall x)(r(x, x) \Rightarrow k(q))$. But this is $f(p)$, and the proof is complete. □

Exercise 7.7. Suppose $\mathscr{R}$ contains only unary relation symbols. Prove that $\mathrm{Pred}(V, \mathscr{R})$ is decidable. (If $p \in P(V, \mathscr{R})$ involves n distinct relation symbols, show that $\models p$ if and only if p is true in every interpretation in a set of at most 2^n elements. This can be done by taking any interpretation M, putting $m_1 \equiv m_2$ if $v(\rho(m_1)) = v(\rho(m_2))$ for all relevant ρ, and working with the equivalence classes.)

Chapter X

Hilbert's Tenth Problem, Word Problems

§1 Hilbert's Tenth Problem

A recursive function $f:\mathbf{N}^n \to \mathbf{N}$ has been defined as one for which there is a Turing machine, T_f say, which computes $f(x_1, \ldots, x_n)$ for all $(x_1, \ldots, x_n) \in \mathbf{N}^n$. Accordingly, in order to show that a particular function $g:\mathbf{N}^n \to \mathbf{N}$ is recursive, we must construct a Turing machine which computes g. This is a tiresome process, even for functions of relatively simple form, and consequently it is natural to seek an alternative characterisation of recursive functions that will facilitate their recognition.

We are accustomed to constructing or decomposing complicated functions in terms of simple functions in other branches of mathematics—for example, use of the chain rule in the differential calculus depends upon the possibility of expressing a function as a composition of simpler functions. We therefore ask if it is possible to build up the set of recursive functions by starting with a set of simple functions and applying certain permissible operations to them. The fact that this can be done is remarkable, for not only does it provide an algebraic characterisation of recursive functions, but it also offers strong support for a belief (known as "Church's Thesis") that all formulations of the concept of an "effectively computable" function must be equivalent (i.e., must produce the same set of functions). For a detailed proof of the characterisation given in Theorem 1.2 below, we refer the reader to [3]. Other accounts of the subject may be found, for example, in [6], [10] or [13].

As initial functions, we take the set I consisting of

(i) the zero function $c: \mathbf{N} \to \mathbf{N}$ given by $c(x) = 0$,

(ii) the successor function $s: \mathbf{N} \to \mathbf{N}$ given by $s(x) = x + 1$,

(iii) the projection functions $U_i^n:\mathbf{N}^n \to \mathbf{N}$ ($n \in \mathbf{N}^+$, $i = 1, \ldots, n$) given by $U_i^n(x_1, \ldots, x_n) = x_i$.

Exercise 1.1. *Show that every initial function is a recursive function.*

The permitted operations are the following:

(i) composition: given $f_j:\mathbf{N}^m \to \mathbf{N}$ and $g:\mathbf{N}^n \to \mathbf{N}$ ($m, n \in \mathbf{N}^+, j = 1, \ldots, n$), composition yields the function $h:\mathbf{N}^m \to \mathbf{N}$ defined by

$$h(x_1, \ldots, x_m) = g(f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m)).$$

(ii) primitive recursion: given $f:\mathbf{N}^n \to \mathbf{N}$ and $g:\mathbf{N}^{n+2} \to \mathbf{N}$ ($n \in \mathbf{N}$), primitive recursion yields the function $h:\mathbf{N}^{n+1} \to \mathbf{N}$ given by

$$h(x_1, \ldots, x_n, 0) = f(x_1, \ldots, x_n),$$
$$h(x_1, \ldots, x_n, t + 1) = g(t, h(x_1, \ldots, x_n, t), x_1, \ldots, x_n) \ (t \in \mathbf{N}).$$

(iii) minimalisation: given $f, g:\mathrm{N}^{n+1} \to \mathrm{N}$ ($n \in \mathrm{N}^+$) satisfying the condition that for each $(x_1, \ldots, x_n) \in \mathrm{N}^n$ there exists at least one y such that $f(x_1, \ldots, x_n, y) = g(x_1, \ldots, x_n, y)$, minimalisation yields the function $h:\mathrm{N}^n \to \mathrm{N}$ given by

$$\begin{aligned} h(x_1, \ldots, x_n) &= \min_y(f(x_1, \ldots, x_n, y) = g(x_1, \ldots, x_n, y)) \\ &= \text{the least } y \in \mathrm{N} \text{ such that } f(x_1, \ldots, x_n, y) \\ &\qquad\qquad\qquad\qquad = g(x_1, \ldots, x_n, y). \end{aligned}$$

Theorem 1.2. *The set of recursive functions coincides with the set of functions obtainable from the set I of initial functions by finite interations of the above operations.*

Exercises

1.3. Prove that the functions $+(x, y) = x + y$, $\times(x, y) = xy$ and $c_k(x) = k$ ($k \in \mathrm{N}$) are recursive, by using Theorem 1.2. Deduce that every polynomial $P:\mathrm{N}^n \to \mathrm{N}$ with coefficients in N is a recursive function.

1.4. (Cf Lemma 6.5 of Chapter IX.) Define the pairing function $p:\mathrm{N}^2 \to \mathrm{N}$ by $p(x, y) = \sum_{r=0}^{x+y} r + y$. Prove that p is bijective, and hence show that the functions $\ell, r: \mathrm{N} \to \mathrm{N}$ given by $p(\ell(z), r(z)) = z$ are well-defined. Show that p, ℓ and r are recursive.

Write $z = p(x, y)$, and define the sequence number function $S:\mathrm{N}^2 \to \mathrm{N}$ by the rule that $S(z, i)$ is the least remainder on division of x by $1 + (i + 1)y$. Prove that S is recursive.

Hilbert's tenth problem seeks an algorithm which will determine whether or not an arbitrary polynomial equation with integral coefficients and in any number of variables has a solution in integers. In 1970, Matiyasevich provided the last step in an argument which proves that no such algorithm exists. We shall outline a method of proof given in full detail in a recent expository article [2] by Davis, which also contains a brief historical account and references.

By a polynomial $P = P(x_1, \ldots, x_n)$ we shall mean a polynomial with integral coefficients. By a solution to the diophantine equation $P(x_1, \ldots, x_n) = 0$ we mean a solution in integers $x_1, \ldots, x_n$. Since every $x \in \mathrm{N}$ is expressible as a sum of four squares of elements of N, the existence of an algorithm to test for solutions implies the existence of an algorithm to test for non-negative solutions, for by testing $P(s_1^2 + t_1^2 + u_1^2 + v_1^2, \ldots, s_n^2 + t_n^2 + u_n^2 + v_n^2) = 0$ for solutions, we have tested $P(x_1, \ldots, x_n) = 0$ for non-negative solutions. Therefore we may restrict all variables to the set N, and prove there does not exist an algorithm to test for solutions in N. We interpret "algorithm" as meaning "Turing algorithm", i.e., a procedure that can be carried out by a suitably designed Turing machine. Since we have information about the set of Turing computable (i.e., recursive) functions, we shall try to relate this set to sets defined in terms of solubility criteria for polynomial equations.

Given a polynomial $P(x_1, \ldots, x_n)$, an obvious subset of $\mathbf{N}^n$ related to it is its solution set $S = \{(x_1, \ldots, x_n) | P(x_1, \ldots, x_n) = 0\}$. For $k = 1, \ldots, n - 1$, the projection S_k of S onto the first k coordinates is given by the set of $(x_1, \ldots, x_k)$ such that there exist $x_{k+1}, \ldots, x_n$ for which $P(x_1, \ldots, x_n) = 0$. Thus membership of the set S_k is related directly to the existence of a solution to P. The following definition generalises this relation.

Definition 1.5. (i) $S \subseteq \mathbf{N}^n$ is *diophantine* if there is a polynomial $P(x_1, \ldots, x_n, y_1, \ldots, y_m)$ in $m + n \geqslant n$ variables such that $(x_1, \ldots, x_n) \in S$ if and only if there exist values $y_1, \ldots, y_m$ for which $P(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0$.

(ii) A relation ρ on $\mathbf{N}^n$ is *diophantine* if the set $\{(x_1, \ldots, x_n) | \rho(x_1, \ldots, x_n)$ is true$\}$ is diophantine. In particular, a function $f : \mathbf{N}^n \to \mathbf{N}$ is diophantine if $\{(x_1, \ldots, x_n, y) | y = f(x_1, \ldots, x_n)\}$ is diophantine.

For brevity, we shall write the condition that S is diophantine informally as

$$(x_1, \ldots, x_n) \in S \text{ iff } (\exists y_1, \ldots, y_m)(P(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0).$$

Example 1.6. The subset S of $\mathbf{N}$, consisting of integers which are not powers of 2, is diophantine, because

$$x \in S \text{ iff } (\exists y, z)(x - y(2z + 1) = 0).$$

Exercises

1.7. Show that the composite elements of $\mathbf{N}$ form a diophantine set.

1.8. Prove that the ordering relations $\{(x, y) | x < y\}$ and $\{(x, y) | x \leqslant y\}$ are diophantine relations on $\mathbf{N}^2$.

1.9. Prove that the divisibility relation $\{(x, y) | x$ divides $y\}$ is diophantine.

1.10. Show that the functions $c(x) = 0, s(x) = x + 1$, and $U_i^n(x_1, \ldots, x_n) = x_i$ $(i = 1, \ldots, n)$, are all diophantine.

1.11. $P, Q : \mathbf{N}^n \to \mathbf{N}$ are polynomials, with solution sets S, Trespectively. Show that $S \cap T$, $S \cup T$ are the solution sets of $P^2 + Q^2 = 0$, $PQ = 0$ respectively. Deduce that diophantine sets are closed under finite unions and intersections.

1.12. Show that the functions p, ℓ, r, defined in Exercise 1.4, are diophantine, and then use Exercise 1.11 to show that the sequence number function $S(z, i)$ is also diophantine.

We have found (Exercise 1.3) that every polynomial function with coefficients in $\mathbf{N}$ is recursive. This result extends to diophantine functions.

Lemma 1.13. *Every diophantine function f is recursive.*

Proof. Write

$$y = f(x_1, \ldots, x_n) \text{ iff } (\exists t_1, \ldots, t_m)(P(x_1, \ldots, x_n, y, t_1, \ldots, t_m) = Q(x_1, \ldots, x_n, y, t_1, \ldots, t_m)),$$

where P, Q are polynomials with coefficients in N. Denoting the sequence number function by $S(z, i)$, then Lemma 6.5 of Chapter IX shows that there exists, for every choice of $y, t_1, \ldots, t_m$, a value u such that $S(u, 0) = y, S(u, 1) = t_1, \ldots, S(u, m) = t_m$. Since f is a function, there is exactly one y for which $P = Q$, hence

$$f(x_1, \ldots, x_n) = y = S\big(\min_u(P(x_1, \ldots, x_n, S(u, 0), \ldots, S(u, m)) \\ = Q(x_1, \ldots, x_n, S(u, 0), \ldots, S(u, m))), 0\big),$$

which, by Exercise 1.4 and Theorem 1.2, shows that f is recursive. □

The essential difficulties arise in attempting to prove the converse to the above result. Using Theorem 1.2, it suffices to prove that every initial function is diophantine, and that the diophantine functions are closed with respect to the operations of composition, primitive recursion and minimalisation. Some of this is easy. Exercise 1.10 has dealt with the initial functions, while if $f_1, \ldots, f_n$ and g are diophantine, and if $h(x_1, \ldots, x_m) = g(f_1(x_1, \ldots, x_m), \ldots, f_n(x_1, \ldots, x_m))$, then so is h, because

$$y = h(x_1, \ldots, x_m) \text{ iff } (\exists t_1, \ldots, t_n)(t_1 = f_1(x_1, \ldots, x_m) \text{ and} \ldots \text{and} \\ t_n = f_n(x_1, \ldots, x_m) \text{ and } y = f(t_1, \ldots, t_n)),$$

which, by Exercise 1.11, is sufficient to establish the result. So it remains to deal with the operations of primitive recursion and minimalisation, neither of which has yet been shown to be expressible in terms of operations which trivially preserve the property of being diophantine. Each of these operations is expressible in terms of the operation of bounded universal quantification, which is now known to preserve this property. A bounded universal quantifier is one which applies for those values of the quantified variable which are less than a given bound. We use the notation $(\forall y \leqslant x)(\ldots)$ to mean "for all $y \in$ N, either $y > x$ or $(\ldots)$". The next theorem is proved in full in [2].

Theorem 1.14. *Let P:N^{m+n+2} → N be a polynomial. Then*

$$S = \{(y, x_1, \ldots, x_n) | (\forall z \leqslant y)((\exists y_1, \ldots, y_m) \\ (P(y, z, x_1, \ldots x_n, y_1, \ldots, y_m) = 0))\}$$

is diophantine.

Corollary 1.15. *The set of diophantine functions is closed under primitive recursion and minimalisation.*

Proof of the Corollary. Suppose f, g are diophantine, and

$$h(x_1, \ldots, x_n, 0) = f(x_1, \ldots, x_n),$$
$$h(x_1, \ldots, x_n, t + 1) = g(t, h(x_1, \ldots, x_n, t), x_1, \ldots, x_n).$$

Using the sequence number function to represent the numbers $h(x_1, \ldots, x_n, 0)$, $\ldots, h(x_1, \ldots, x_n, z)$, we have $y = h(x_1, \ldots, x_n, z)$ if and only if

$$(\exists u)\big((\exists v)(v = S(u, 0) \wedge v = f(x_1, \ldots, x_n)) \wedge (\forall t \leqslant z)\big(t = z \vee (\exists w)(w = S(u, t + 1) \wedge w = g(t, S(u, t), x_1, \ldots, x_n))\big) \wedge y = S(u, z)\big)$$

which, by Exercises 1.11 and 1.12, shows that h is diophantine.

Finally, if f, g are diophantine and

$$h(x_1, \ldots, x_n) = \min_y(f(x_1, \ldots, x_n, y) = g(x_1, \ldots, x_n, y)),$$

then $y = h(x_1, \ldots, x_n)$ if and only if

$$(\exists z)(z = f(x_1, \ldots, x_n, y) \wedge z = g(x_1, \ldots, x_n, y)) \wedge (\forall t \leqslant y)(t = y \vee (\exists u)(\exists v)(u = f(x_1, \ldots, x_n, t) \wedge v = g(x_1, \ldots, x_n, t) \wedge (u < v \vee v < u)))$$

showing that h is diophantine. □

We may therefore state the following fundamental result.

Theorem 1.16. *A function is recursive if and only if it is diophantine.*

In chapter IX, we showed the existence of a subset E of N which is recursively enumerable but not recursive. That is, E is the range of some recursive function, but the characteristic function of E is not a recursive function. Theorem 1.16 implies that a subset of N is recursively enumerable if and only if it is diophantine. Hence E is diophantine, and so there is a polynomial P such that

$$x \in E \text{ iff } (\exists t_1, \ldots, t_m)(P(x, t_1, \ldots, t_m) = 0).$$

Suppose that there exists a Turing machine M which can test every polynomial equation for the existence of solutions. M, when applied to the sequence of polynomials $P(0, t_1, \ldots, t_m)$, $P(1, t_1, \ldots, t_m)$, ..., will then compute the characteristic function of E. Thus E has a recursive characteristic function and hence is a recursive set, which contradicts its definition. Therefore, no such Turing machine M can exist. This statement is to be considered as an explicit denial of the existence of any algorithm to test all polynomial diophantine equations for solutions, which therefore implies that Hilbert's tenth problem is insoluble.

Exercises

1.17. Prove that a subset of N is recursively enumerable if and only if it is diophantine.

1.18. Give an enumeration of the set of polynomials with integral coefficients and in an arbitrary finite number of variables chosen from $x, y_1, y_2, \ldots$. Hence obtain a sequence $\{D_n\}$ which contains all diophantine subsets of N. Define a function $g{:}\mathrm{N}^2 \to \mathrm{N}$ by

$$\begin{aligned} g(x, n) &= 0 \quad \text{if} \quad x \notin D_n, \\ g(x, n) &= 1 \quad \text{if} \quad x \in D_n. \end{aligned}$$

Use Theorem 1.16 to prove that g is not recursive. Obtain an alternative proof that Hilbert's tenth problem is insoluble by showing that the existence of a "Hilbert algorithm" would imply that g is recursive.

§2 Word Problems

A group G is often specified by giving a set X of generators of G together with a set R of relations satisfied by these generators. The set R is required to be such that every relation on the elements of X which holds in G is a consequence of those in R. Here, a relation is an equation $w_1(a_1, \ldots, a_n) = w_2(a_1, \ldots, a_n)$ which holds in G, where $a_1, \ldots, a_n$ are particular elements of X and w_1, w_2 are group theoretical words. We can express such an equation in the form $w_1(a_1, \ldots, a_n)(w_2(a_1, \ldots, a_n))^{-1} = 1$, so we may always suppose that each relation is given in the form $w(a_1, \ldots, a_n) = 1$, and identify the relation with the word $w(a_1, \ldots, a_n)$.

Definition 2.1. A *group presentation* is a set X together with a set R of group theoretical words on the elements of X. The presentation (X, R) is called *finite* if both X and R are finite.

Every group presentation (X, R) does determine a group: take the free group F on X and the smallest normal subgroup K of F which contains R, and then the group determined by (X, R) is the factor group F/K. We shall write $G = \langle X | R \rangle$ to indicate that G is the group determined by the presentation (X, R). (The group G has of course many different presentations.) Henceforth, in order to avoid confusion between an element of G and a particular construction of the element, a word w shall mean an element of the free group F. The corresponding element of $G = F/K$ will be called the group element represented by w. Two words w_1, w_2 will be called equivalent, written $w_1 \sim w_2$, if they represent the same group element.

The properties of the group $G = \langle X | R \rangle$ may not be apparent from the presentation. From the information in a given presentation of a group, we may be able to obtain answers to various questions about the group, and we are interested in finding procedures for doing this. M. Dehn in 1911 formulated three basic decision problems for a given presentation of a group $G = \langle X | R \rangle$. These three problems are known as the Word Problem, the Conjugacy Problem and the Isomorphism Problem.

Problem 2.2. (The Word Problem) *Find an algorithm which, for each word w in the elements of X, determines whether or not w represents the identity element of G.*

Problem 2.3. (The Conjugacy Problem) *Find an algorithm which, for any two words w_1, w_2, determines whether or not w_1 and w_2 represent conjugate elements of G.*

Problem 2.4. (The Isomorphism Problem) *Find an algorithm which, for any group presentation (X', R'), determines whether or not $\langle X' | R' \rangle$ is isomorphic to G.*

These problems have been solved for certain suitably restricted classes of presentations. (The reader is referred to [9], Section 6.1, for details.) In general, however, these problems are insoluble, and we shall try to show in this section how the theory of Turing machines can be used to establish the insolubility. In order that the underlying ideas will not be obscured by details, we shall restrict ourselves to a demonstration that there is a finitely-presented semigroup S whose word problem is insoluble. The interested reader will find in Chapter 12 of [11] an account of the construction from S of a finitely presented group G with insoluble word problem. (This construction is purely algebraic, and makes no further use of the theory of Turing machines.)

Exercises

2.5. A presentation (X, R) is called *abelian* if, for every $x, y \in X$, we have $x^{-1}y^{-1}xy \in R$. Show that the word problem for a finite abelian presentation is soluble. Show also that the isomorphism problem is soluble for pairs of finite abelian presentations.

2.6. Given that the finitely presented group $G = \langle X|R\rangle$ is finite, prove that it has soluble word problem and soluble conjugacy problem.

We now show how to associate a finite semigroup presentation with a Turing machine M. The idea behind the construction is to regard instantaneous descriptions as words, and to introduce relations which will make an instantaneous description represent the same semigroup element as does the instantaneous description obtained from the former one by one operation of the machine.

We shall always work with the shortest description, thereby avoiding difficulties arising from different descriptions of the same state of M. Thus an instantaneous description shall neither begin nor end with s_0. However, this introduces some difficulty into the construction of the set of relations, which we resolve by use of an end symbol e. With the description $\sigma q_i \tau$, we shall associate the semigroup word $e\sigma q_i \tau e$. It is also convenient to introduce a new internal state symbol q_∞, meaning that the machine has stopped.

As we are dealing with semigroups and not groups, a relation necessarily involves two words, and has the form $w_1 \sim w_2$. For our purposes, it is convenient to regard this as an ordered pair of words, and so to treat $w_1 \sim w_2$ and $w_2 \sim w_1$ as different relations. Each relation then has a first word.

Let M be a Turing machine with alphabet $\mathfrak{S} = \{s_0, s_1, \dots, s_m\}$ and set of internal states $\mathfrak{Q} = \{q_0, q_1, \dots, q_n\}$. The *semigroup presentation associated with* M is the presentation with generator set $X = \mathfrak{S} \cup \mathfrak{Q} \cup \{e, q_\infty\}$ and with relation set R consisting of

(a) for each quadruple $q_i s_j s_k q_\ell \in M$, the relation

$$q_i s_j \sim q_\ell s_k,$$

(b) for each quadruple $q_i s_j L q_\ell \in M$ with $j \neq 0$, the relations

$$s_k q_i s_j \sim q_\ell s_k s_j \qquad (\text{all } k),$$
$$e q_i s_j \sim e q_\ell s_0 s_j,$$

(b_0) for each quadruple $q_is_0Lq_\ell \in M$, the relations

$$\begin{aligned} s_kq_is_0 &\sim q_\ell s_ks_0 \qquad (\text{all } k),\\ eq_is_0 &\sim eq_\ell s_0s_0,\\ s_kq_ie &\sim q_\ell s_ke \qquad (\text{all } k \neq 0)\\ s_0q_ie &\sim q_\ell e,\\ eq_ie &\sim eq_\ell e, \end{aligned}$$

(c) for each quadruple $q_is_jRq_\ell \in M$ with $j \neq 0$, the relation

$$q_is_j \sim s_jq_\ell,$$

(c_0) for each quadruple $q_is_0Rq_\ell \in M$, the relations

$$\begin{aligned} s_kq_is_0 &\sim s_ks_0q_\ell \qquad (\text{all } k),\\ eq_is_0 &\sim eq_\ell,\\ s_kq_ie &\sim s_ks_0q_\ell e \quad (\text{all } k),\\ eq_ie &\sim eq_\ell e, \end{aligned}$$

(d) for each pair q_is_j for which there is no quadruple in M beginning with q_is_j, the relation

$$q_is_j \sim q_\infty s_j,$$

and, if $j = 0$, the relation

$$q_ie \sim q_\infty e.$$

Let $w_1 = \sigma a\tau$ be a word and let $a \sim b$ be a relation in the above list. Substitution of b for a in w_1 gives the equivalent word $w_2 = \sigma b\tau$. Such a substitution, where the second member of a relation is substituted for the first, will be called a *forward step*. We write $w_1 \rightarrow w_2$ to denote that w_2 is obtainable from w_1 by a forward step. The reverse substitution is called a *backward step*, and we write $w_2 \leftarrow w_1$ to denote that w_1 is obtainable from w_2 by a backward step. We write $w - w'$ to denote that w' is obtainable from w by a step which may be either forward or backward. A *path* from w to w' is a finite sequence of steps $w - w_1 - w_2 - \cdots - w_{n-1} - w'$ beginning with w and ending with w'. Clearly, two words w, w' are equivalent if and only if there exists a path from w to w'.

We now concentrate our attention on the words which correspond to an instantaneous description of the Turing machine.

Definition 2.7. A *special word* on X is a word of the form $e\sigma q_i\tau e$, where σ, τ are words (possibly empty) on $\mathfrak{S}$, such that σ does not begin with s_0 and τ does not end with s_0. The special word $e\sigma q_i\tau e$ is called *terminal* if $i = \infty$.

Any word obtained from a special word by a step is again a special word. Forward steps on special words correspond to steps in the operation of the machine M.

Lemma 2.8. *Let w, w' be special words. Suppose w' is terminal. Then w, w' are equivalent if and only if there is a path from w to w' consisting only of forward steps.*

Proof. Trivially, if such a path exists, then $w \sim w'$. Suppose that $w \sim w'$. Then there exists a path

$$w = w_0 - w_1 - \cdots - w_n = w'$$

from w to w'. We may suppose the path is chosen so that the number n of steps is the least possible. (If $n = 0$, then the path consists only of forward steps.) If the path has any backward steps, then there is a last such, say $w_k \leftarrow w_{k+1}$. This cannot be the last step of the path, because there is no forward step away from a terminal word. Thus $k + 1 < n$ and $w_{k+1} \rightarrow w_{k+2}$ is a forward step. But there is at most one forward step away from any special word, since the machine operation is determined. This implies that $w_{k+2} = w_k$, and so

$$w = w_0 - w_1 - \cdots - w_k - w_{k+3} - \cdots - w_n = w'$$

is a shorter path from w to w', contrary to the original choice of path. Hence the shortest path consists only of forward steps. □

We are now able to produce a Turing machine whose associated semigroup presentation has insoluble word problem.

Theorem 2.9. *Let E be a recursively enumerable but non-recursive subset of* $\mathbf{N}$, *and let M be a Turing machine which, when started in the state $q_0 s_1^n$, stops with blank tape if $n \in E$, and does not stop if $n \notin E$. Then the semigroup presentation associated with M has insoluble word problem.*

Remark. The existence of such a set E and Turing machine M was established in Exercises 5.6 and 5.8 of Chapter IX.

Proof. By Lemma 2.8, the special word $eq_0 s_1^n e$ is equivalent to $eq_\infty e$ if and only if there exists a forward path from $eq_0 s_1^n e$ to $eq_\infty e$. Such a path exists if and only if M, started in the state $q_0 s_1^n$, stops with blank tape—i.e., if and only if $n \in E$. Since E is non-recursive, the word problem (even for this restricted set of words) is recursively insoluble. □

References and Further Reading

[1] Bell, J. L., Slomson, A. B.: *Models and Ultraproducts: An Introduction.* First revised edition. Amsterdam: North-Holland 1971.

[2] Davis, Martin: Hilbert's Tenth Problem is Unsolvable. *Amer. Math. Monthly* **80**, 233–269 (1973).

[3] Davis, Martin: *Computability and Unsolvability.* New York: McGraw-Hill 1958.

[4] Halmos, P. R.: *Naive Set Theory.* Princeton: Van Nostrand 1960. New York–Heidelberg–Berlin: Springer 1974.

[5] Hermite, Charles: *Oeuvres, tII,* pp. 5–12 Paris: Gauthier-Villars 1908.

[6] Kleene, S. C.: *Introduction to Metamathematics.* Princeton: Van Nostrand 1952.

[7] Lyndon, Roger C: *Notes on Logic. Mathematical Studies,* no. 6. New York: Van Nostrand 1964.

[8] Machover, M., Hirschfeld, J.: *Lectures on Non-Standard Analysis. Lecture Notes in Mathematics* 94. Berlin–Heidelberg–New York: Springer 1969.

[9] Magnus, W., Karrass, A., Solitar, D.: *Combinatorial Group Theory: Presentations of Groups in Terms of Generators and Relations.* New York–London–Sydney: Wiley-Interscience 1966.

[10] Mendelson, E.: *Introduction to Mathematical Logic.* New York: Van Nostrand 1964.

[11] Rotman, Joseph J.: *The Theory of Groups: An Introduction.* Boston- Allyn and Bacon 1965.

[12] Stoll, Robert R.: *Set Theory and Logic.* San Francisco-London: Freeman 1963.

[13] Yasuhara, Ann: *Recursive Function Theory and Logic.* New York: Academic Press 1971.

[14] Yates, Robert C.: *The Trisection Problem.* Reston: The National Council of Teachers of Mathematics 1971.

Index of Notations

The following notations are used at points remote from their explanations, which are given on the pages indicated.

$A \models p$	p is a consequence of the assumptions A	13
$A \vdash p$	p is provable from the assumptions A	15
$\mathrm{al}(n)$	formula expressing "at least n elements"	45
$\mathrm{ar}(t)$	arity of t	2
$\mathrm{Con}(A)$	set of consequences of A	13
$d(p)$	depth of quantification of p	28
$D\rho$	domain of the relation ρ	76
$\mathrm{Ded}(A)$	set of all deductions from A	15
$\mathrm{dist}(x_1, \ldots, x_n)$	formula expressing "$x_1, \ldots, x_n$ are distinct"	48
$G(p)$	Gödel number of p	91, 92
$\mathcal{I}$	identity relation symbol	38
I	(when unexplained) set of axioms of identity	38
$\mathcal{L}(\mathcal{T})$	language of the theory $\mathcal{T}$	40
$\mathcal{N}$	recursive arithmetic	96
$\mathcal{N}_0$	recursive arithmetic without induction axiom scheme	97
$\mathcal{O}^k(S)$	set of kth order objects on S	75
$\mathcal{P}$	Peano arithmetic	59
$P(\mathcal{T})$	algebra of the theory $\mathcal{T}$	40
$P(X)$	free proposition algebra on X	12
$\tilde{P}(V, \mathcal{R})$	full first-order algebra on $(V, \mathcal{R})$	27
$P(V, \mathcal{R})$	reduced first-order algebra on $(V, \mathcal{R})$	28
$\mathrm{Pow}(M)$	power set of M	54
$\mathrm{Pred}(V, \mathcal{R})$	first-order predicate calculus on $(V, \mathcal{R})$	30
$\mathrm{Pred}_{\mathcal{I}}(V, \mathcal{R})$	first-order predicate calculus with identity, on $(V, \mathcal{R})$	39
$\mathrm{Prop}(X)$	propositional calculus on X	18
$\Psi_M^{(k,\ell)}$	partial recursive function $\mathrm{N}^k \to \mathrm{N}^\ell$ defined by Turing machine M	89
$\mathcal{R}$	set of relation symbols	27
$\mathcal{R}_n$	subset of n-ary relation symbols	27
$\mathrm{rel}(M)$	set of all relations of all arities on M	42
S^σ	enlargement of S with respect to σ	77
$^*\mathsf{S}$	full enlargement of S	77
$\mathcal{T}(\mathsf{S})$	theory of the system S	74
$v(p)$	truth value of p	13
$\mathrm{var}(p)$	set of (free) variables of p	29

Subject Index

Addition, theory of 61
Adequacy 18
Adequacy Theorem 22, 36
Algebra, free 4
  full first-order 27
  Lindenbaum 22
  proposition 12
  reduced first-order 28
  relatively free 8
  universal 2
Algebraic theory 72
Alphabet of Turing machine 85
  universal 90
Arity 1
Axiom of Choice 55, 57
  of Extension 53, 56
  of Infinity 54, 57
  of Pairing 53, 57
  of Power Set 54, 57
  of Regularity 56
  of Union 54, 57
Axiom Schema of Replacement 55, 57
  of Restriction 55, 57
  of Subsets 53, 56
Axioms of Pred(V, $\mathscr{R}$) 30
  of Prop(X) 15
  of a theory 40

Bounded universal quantifier 108

Cardinal of model 44
  of theory 45
Categorical 44, 47
  in cardinal χ 45, 47
Church's Theorem 102
Church's Thesis 105
Compactness Theorem 22, 37
Completeness 43, 44, 47, 51
Computation by Turing machine 88
Concurrent relation 76
Conjugacy Problem 110
Consequence 13, 30
  proper 39
Consistency 18
  of first-order theory 43
Consistency Theorem 19, 34
Consistent subset 40
  maximal 40
Constant 40
Continuity 83
  uniform 83

Decidability 19, 94
  of Prop(X) 24
  of Theory of Equality 51
Deduction 15, 31
Deduction Theorem 19, 34
Definable 43
  strongly 97
Dense Linear Order 47
Depth of quantification 28
Description, of relation 76
  instantaneous, of Turing machine 87
  shortest 87
Diophantine set 107
Domain of relation 76
Dual 42

Effectively axiomatised theory 93
Elementary 41
Elementary Group Theory 41
Elementary Theory of Fields 45
Enlargement of theory 77
  of standard object 77
  full 77
Extension of model 77
  of theory 77

Failure of Turing machine 88
Field, algebraic closure of 66
Fields, Elementary Theory of 45
Filter 63
  existence of 65
  Fréchet 63
Finite intersection property 65

Generalisation 31
Gödel number 91, 92
Gödel's Completeness Theorem 37
  see Adequacy Theorem 36
Gödel's Incompleteness Theorem 100

Hilbert's Tenth Problem 105
Homomorphism 13

Identical relation 7
Identity, axioms of 38
  Predicate Calculus with 39
Incompleteness of
    recursive arithmetic 100
Infinite real numbers 81
Infinitesimal real numbers 81
Initial state 88
Internal relation 80
Interpretation 30
  proper 39
Isomorphism of models 44
  of *T*-algebras 4
Isomorphism Problem 110

Kalmar's Theorem 103
*k*th-order objects 75

Language of theory 40
Law of an algebra 7
  of a variety 7
Logic 18
Limit, direct 70
    existence of 72
  of function 83
  of sequence 82
Löwenheim-Skolem Theorem 46

Minimalisation 106
Model of first-order theory 42
  of Peano arithmetic 61
  non-standard 75
  standard 75
Modus ponens 15
Monad 82

Natural numbers in ZF 54
Non-standard model 75
Normal form, conjunctive 24
  disjunctive 24
  prenex 37

Operation 1, 2

$\mathscr{P}$ (Peano arithmetic) 59
  completeness of 61
  models of 61
Pair 53
Pairing function 106
Path 112
Peano axioms 58, 59
Predicate 26
Presentation 110, 111
Primary proposition 49
Primitive recursion 105
Projective geometry 40
Problem, Conjugacy 110
  Decision 94
  Hilbert's Tenth 105
  Isomorphism 110
  Stopping 95
  Word 110
Proof 15, 31
Proper consequence 39
  interpretation 39

Quantifier, bounded universal 108
  elimination 48
  existential 26
  reduction 48
  universal 26

Recursive arithmetic 97
  function 89, 106
  subset 90
Recursively enumerable 90, 94
  but non-recursive 95
Recursively soluble 94
Relation, concurrent 76
  definable 43
  external 80
  identical 7
  internal 80
Russell Paradox 53

Satisfiability Theorem 21, 36, 39
Semantic implication 13, 30
Semigroup of Turing machine 111
Semi-homomorphism 32, 33
Sequence number function 97
Soundness 18
Soundness Theorem 19, 33
Special word 112
Standard model 75
Standard relation 78
Standard system 74
  higher-order 76
  language of 74
  theory of 74
State of Turing machine 87
  initial 88
  terminal 88

State function 97
Step 112
Stopping Problem 95
Strongly definable 97
Substitution Theorem 17, 33
Subsystem of standard system 77
Subultraproduct 69
Successor 54
Syntactic implication 15, 31

T-algebra 2
T-subalgebra 3
Tautology 14, 30
Terminal state 88
  word 112
Theorem 15, 31, 40, 100
Theory 40
Transitive set 55
Truth 13, 29
  function 23
Turing machine 85
  universal 92
Type 2

Ultrafilter 63
  α-complete 67
  existence of 65
  principal 63
  restriction of 63
  uniform 65
Ultrapowers 67
Ultraproducts 64
Undecidability of
    recursive arithmetic 100
  of Predicate Calculus 101
Uniform convergence 84
Uniform ultrafilter 65
Universal algebra 1
  alphabet 90
  proposition 72
  Turing machine 92

Valid 14, 30
Valuation 13
Variable 6
Variables involved in element of $\tilde{P}$ 27
  of element of P 29
Variety of *T*-algebras 7

Word 6
  special 112
Word Problem 110