Introduction to Mathematical Logic

SECOND EDITION

ELLIOTT MENDELSON

Introduction to Mathematical Logic

SECOND EDITION

ELLIOTT MENDELSON

Queens College of the City University of New York

D. VAN NOSTRAND COMPANY

New York Cincinnati Toronto London Melbourne

To Arlene

D. Van Nostrand Company Regional Offices:
New York Cincinnati

D. Van Nostrand Company International Offices:
London Toronto Melbourne

Library of Congress Catalog Card Number: 78-65959
ISBN: 0-442-25307-9

Published by D. Van Nostrand Company
135 West 50th Street, New York, N.Y. 10020

10 9 8 7 6 5 4 3 2 1

PREFACE TO THE SECOND EDITION

This new edition contains considerable improvements over the first edition. Much new material has been added. For example, in Chapter 2 there are two new sections on model theory devoted to elementary equivalence and elementary extensions and to ultrapowers and nonstandard analysis. The greatest change has been the addition of a large number of exercises. There are 389 exercises, many of them consisting of several parts. Completely new is a section at the end of the book, Answers to Selected Exercises, which should improve the usefulness of the book as a textbook as well as for independent study. With all these changes, I have attempted to preserve the spirit of the original book, which was intended to be a simple, clear introduction to mathematical logic unencumbered by excessive notation and terminology.

I should like to thank the many people who have given me suggestions for corrections and improvement. I am particularly indebted to Professor Frank Cannonito for much helpful advice.

ELLIOTT MENDELSON

PREFACE TO THE FIRST EDITION

In this book we have attempted to present a compact introduction to some of the principal topics of mathematical logic. In order to give a full and precise treatment of the more important basic subjects, certain subsidiary topics, such as modal, combinatory, and intuitionistic logics, and some interesting advanced topics, such as degrees of recursive unsolvability, have had to be omitted.

In the belief that beginners should be exposed to the most natural and easiest proofs, free-swinging set-theoretic methods have been used. The significance of a demand for constructive proofs can be evaluated only after a certain amount of experience with mathematical logic has been obtained. After all, if we are to be expelled from "Cantor's paradise" (as non-constructive set theory was called by Hilbert), at least we should know what we are missing.

The five chapters of the book can be covered in two semesters, but, for a one-semester course, Chapters 1 through 3 will be quite adequate (omitting, if hurried, Sections 5 and 6 of Chapter 1 and Sections 10, 11, and 12 of Chapter 2). The convention has been adopted of prefixing a superscript "D" to any section or exercise which will probably be difficult for a beginner, and a superscript "A" to any section or exercise which presupposes familiarity with a topic that has not been carefully explained in the text. Bibliographical references are given to the best source of information, which is not always the earliest paper; hence these references give no indication as to priority. For example, Boone [1959] gives the most complete account of his work on the word problem, which was actually done independently of and about the same time as Novikov's work [1955].

The present book is an expansion of lecture notes for a one-semester course in mathematical logic given by the author at Columbia University from 1958 to 1960 and at Queens College in 1961 and 1962. The author hopes that it can be read with ease by anyone with a certain amount of experience in abstract mathematical thought, but there is no specific prerequisite. The author would like to thank J. Barkley Rosser for encouragement and guidance during his graduate studies in logic, and he would like to acknowledge also the obvious debt owed to the books of Hilbert-Bernays, 1934, 1939; Kleene, 1952; Rosser, 1953; and Church, 1956.

CONTENTS

		PAGE
	INTRODUCTION	**1**
CHAPTER		
1	**THE PROPOSITIONAL CALCULUS**	**11**
1.	Propositional Connectives. Truth Tables	11
2.	Tautologies	16
3.	Adequate Sets of Connectives	25
4.	An Axiom System for the Propositional Calculus	29
5.	Independence. Many-Valued Logics	38
6.	Other Axiomatizations	40
2	**QUANTIFICATION THEORY**	**45**
1.	Quantifiers	45
2.	Interpretations. Satisfiability and Truth. Models	50
3.	First-Order Theories	58
4.	Properties of First-Order Theories	61
5.	Completeness Theorems	65
6.	Some Additional Metatheorems	72
7.	Rule C	76
8.	First-Order Theories with Equality	79
9.	Definitions of New Function Letters and Individual Constants	85
10.	Prenex Normal Forms	88
11.	Isomorphism of Interpretations. Categoricity of Theories	93
12.	Generalized First-Order Theories. Completeness and Decidability	95
13.	Elementary Equivalence. Elementary Extensions	103
14.	Ultrapowers. Non-Standard Analysis	108

CHAPTER PAGE

3 FORMAL NUMBER THEORY 121

1. An Axiom System 121
2. Number-Theoretic Functions and Relations 134
3. Primitive Recursive and Recursive Functions 137
4. Arithmetization. Gödel Numbers 151
5. Gödel's Theorem for S 158
6. Recursive Undecidability. Tarski's Theorem. Robinson's System 165

4 AXIOMATIC SET THEORY 173

1. An Axiom System 173
2. Ordinal Numbers 183
3. Equinumerosity. Finite and Denumerable Sets 192
4. Hartogs' Theorem. Initial Ordinals. Ordinal Arithmetic 200
5. The Axiom of Choice. The Axiom of Regularity 209

5 EFFECTIVE COMPUTABILITY 221

1. Markov Algorithms 221
2. Turing Algorithms 240
3. Herbrand–Gödel Computability. Recursively Enumerable Sets 248
4. Undecidable Problems 265

BIBLIOGRAPHY 269

NOTATION 289

ANSWERS TO SELECTED EXERCISES 293

INDEX 321

INTRODUCTION

One of the most popular definitions of logic is that it is the analysis of methods of reasoning. In studying these methods, logic is interested in the form rather than the content of the argument. For example, consider the two deductions:

(1) All men are mortal. Socrates is a man. Hence Socrates is mortal.
(2) All rabbits like carrots. Sebastian is a rabbit. Hence, Sebastian likes carrots.

Both have the same form: All A are B. S is an A. Hence S is a B. The truth or falsity of the particular premisses and conclusions is of no concern to the logician. He wants to know only whether the truth of the premisses implies the truth of the conclusion. The systematic formalization and cataloguing of valid methods of reasoning is one of the main tasks of the logician. If his work uses mathematical techniques and if it is primarily devoted to the study of mathematical reasoning, then it may be called mathematical logic. We can narrow the domain of mathematical logic if we define its principal aim to be a precise and adequate definition of the notion of "mathematical proof".

Impeccable definitions have little value at the beginning of the study of a subject. The best way to find out what mathematical logic is about is to start doing it, and the student is advised to begin reading the book even though (or especially if) he has qualms about the meaning or purposes of the subject.

Although logic is basic to all other studies, its fundamental and apparently self-evident character discouraged any deep logical investigations until the late nineteenth century. Then, under the impetus of the discovery of non-Euclidean geometries and of the desire to provide a rigorous foundation for analysis, interest in logic revived. This new interest, however, was still rather unenthusiastic until, around the turn of the century, the mathematical world was shocked by the discovery of the paradoxes, i.e., arguments leading to contradictions. The most important of these paradoxes are the following.

Logical Paradoxes

(1) (Russell, 1902) By a set, we mean any collection of objects, e.g., the set of all even integers, the set of all saxophone players in Brooklyn, etc. The objects which make up a set are called its members. Sets may themselves be members of sets, e.g., the set of all sets of integers has sets as its members. Most sets are not members of themselves; the set of cats, for example, is not a member of itself, because the set of cats is not a cat. However, there may be sets which do belong to themselves, e.g., the set of all sets. Now, consider the set A of all those sets X such that X is not a member of X. Clearly, by definition, A is a member of A if and only if A is not a member of A. So, if A is a member of A, then A is also not a member of A; and if A is not a member of A, then A is a member of A. In any case, A is a member of A and A is not a member of A.

(2) (Cantor, 1899) This paradox involves a certain amount of the theory of cardinal numbers and may be skipped by those having no previous acquaintance with that theory. The cardinal number $\overline{\overline{Y}}$ of a set Y is defined to be the set of all sets X which are equinumerous with Y (i.e., for which there is a one-one correspondence between Y and X, cf. page 7). We define $\overline{\overline{Y}} \leqslant \overline{\overline{Z}}$ to mean that Y is equinumerous with a subset of Z; by $\overline{\overline{Y}} < \overline{\overline{Z}}$ we mean $\overline{\overline{Y}} \leqslant \overline{\overline{Z}}$ and $Y \neq \overline{\overline{Z}}$. Cantor proved that, if $\mathcal{P}(Y)$ is the set of all subsets of Y, then $\overline{\overline{Y}} < \overline{\overline{\mathcal{P}(Y)}}$ (cf. page 195). Let C be the universal set, i.e., the set of all sets. Now, $\mathcal{P}(C)$ is a subset of C, so it follows easily that $\overline{\overline{\mathcal{P}(C)}} \leqslant \overline{\overline{C}}$. On the other hand, by Cantor's Theorem, $\overline{\overline{C}} < \overline{\overline{\mathcal{P}(C)}}$. The Schröder-Bernstein Theorem (cf. page 194) asserts that if $\overline{\overline{Y}} \leqslant \overline{\overline{Z}}$ and $\overline{\overline{Z}} \leqslant \overline{\overline{Y}}$, then $\overline{\overline{Y}} = \overline{\overline{Z}}$. Hence, $\overline{\overline{C}} = \overline{\overline{\mathcal{P}(C)}}$, contradicting $\overline{\overline{C}} < \overline{\overline{\mathcal{P}(C)}}$.

(3) (Burali-Forti, 1897) This paradox is the analogue in the theory of ordinal numbers of Cantor's Paradox and will make sense only to those already familiar with ordinal number theory. Given any ordinal number, there is a still larger ordinal number. But the ordinal number determined by the set of all ordinal numbers is the largest ordinal number.

Semantic Paradoxes

(4) The Liar Paradox. A man says, "I am lying." If he is lying, then what he says is true, and so he is not lying. If he is not lying, then what he says is true, and so he is lying. In any case, he is lying and he is not lying.†

†The Cretan "paradox", known in antiquity, is similar to the Liar Paradox. The Cretan philosopher Epimenides said, "All Cretans are liars." If what he said is true, then, since Epimenides is a Cretan, it must be false. Hence, what he said is false. Thus, there must be some Cretan who is not a liar. This is not logically impossible, so we do not have a genuine paradox. However, the fact that the utterance by Epimenides of that false sentence could imply the existence of some Cretan who is not a liar is rather unsettling.

(5) (Richard, 1905) Some phrases of the English language denote real numbers, e.g., "the ratio between the circumference and diameter of a circle" denotes the number π. All phrases of the English language can be enumerated in a standard way: order all phrases having k letters lexicographically (as in a dictionary), and then place all phrases with k letters before all phrases with a larger number of letters. Hence, all phrases of the English language denoting real numbers can be enumerated merely by omitting all other phrases in the given standard enumeration. Call the n^{th} real number in this enumeration the n^{th} Richard number. Consider the phrase: "the real number whose n^{th} decimal place is 1 if the n^{th} decimal place of the n^{th} Richard number is not 1, and whose n^{th} decimal place is 2 if the n^{th} decimal place of the n^{th} Richard number is 1". This phrase defines a Richard number, say the k^{th} Richard number; but, by its definition, it differs from the k^{th} Richard number in the k^{th} decimal place.

(6) (Berry, 1906) There are only a finite number of syllables in the English language. Hence, there are only a finite number of English expressions containing fewer than forty syllables. There are, therefore, only a finite number of positive integers which are denoted by an English expression containing fewer than forty syllables. Let k be *the least positive integer which is not denoted by an expression in the English language containing fewer than forty syllables*. The italicized English phrase contains fewer than forty syllables and denotes the integer k.

(7) (Grelling, 1908) An adjective is called *autological* if the property denoted by the adjective holds for the adjective itself. An adjective is called *heterological* if the property denoted by the adjective does not apply to the adjective itself. For example, "polysyllabic" and "English" are autological, while "monosyllabic", "French", and "blue" are heterological. Consider the adjective "heterological". If "heterological" is heterological, then it is not heterological. If "heterological" is not heterological, then it is heterological. In any case, "heterological" is both heterological and not heterological.

All of these paradoxes are genuine in the sense that they contain no obvious logical flaws. The logical paradoxes involve only notions from the theory of sets, whereas the semantic paradoxes also make use of concepts like "denote", "true", "adjective", which need not occur within our standard mathematical language. For this reason, the logical paradoxes are a much greater threat to a mathematician's peace of mind than the semantic paradoxes.

Analysis of the paradoxes has led to various proposals for avoiding them. All of these proposals are restrictive in one way or another of the "naive" concepts which enter into the derivation of the paradoxes. Russell noted the self-reference present in all the paradoxes and suggested that every object must have a definite non-negative integer as its "type". Then an expression, "x is a member of the set y", is *meaningful* if and only if the type of y is one greater than the type of x.

This approach, known as the theory of types and systematized and developed by Russell-Whitehead [1910–1913], is successful in eliminating the known paradoxes,† but it is clumsy in practice and has certain other drawbacks as well. A different criticism of the logical paradoxes is aimed at their assumption that, for every property $P(x)$, there exists a corresponding set of all objects x which satisfy $P(x)$. If we reject this assumption, then the logical paradoxes are no longer derivable.‡ It is necessary, however, to provide new postulates that will enable us to prove the existence of those sets which are a daily necessity to the practicing mathematician. The first such axiomatic set theory was invented by Zermelo [1908]. In Chapter 4 we shall present an axiomatic theory of sets which is a descendant of Zermelo's system (with some new twists given to it by von Neumann, R. Robinson, Bernays, and Gödel). There are also various hybrid theories combining some aspects of type theory and axiomatic set theory, e.g., Quine's system NF (cf. Rosser [1953]).

A more radical interpretation of the paradoxes has been advocated by Brouwer and his intuitionist school (cf. Heyting [1956]). They refuse to accept the universality of certain basic logical laws, such as the law of excluded middle: P or not-P. Such a law, they claim, is true for finite sets, but it is invalid to extend it on a wholesale basis to all sets. Likewise, they say it is invalid to conclude that "there exists an object x such that not-$P(x)$" follows from "not-(for all x, $P(x)$)"; we are justified in asserting the existence of an object having a certain property only if we know an effective method for constructing (or finding) such an object. The paradoxes are, of course, not derivable (or even meaningful) if we obey the intuitionist strictures, but, alas, so are many beloved theorems of everyday mathematics, and, for this reason, intuitionism has found few converts among mathematicians.

Whatever approach one takes to the paradoxes, it is necessary first to examine the language of logic and mathematics to see what symbols may be used, to determine the ways in which these symbols are put together to form terms, formulas, sentences, and proofs, and to find out what can and cannot be proved if certain axioms and rules of inference are assumed. This is one of the tasks of mathematical logic, and, until it is done, there is no basis for comparing rival foundations of logic and mathematics. The deep and devastating results of Gödel, Tarski, Church, Rosser, Kleene, and many others have been ample reward for the labor invested and have earned for mathematical logic its status as an independent branch of mathematics.

†Russell's Paradox, for example, depends upon the existence of the set A of all sets which are not members of themselves. Because, according to the theory of types, it is meaningless to say that a set belongs to itself, there can be no such set A.

‡Russell's Paradox then proves that there is no set A of all sets which do not belong to themselves; the paradoxes of Cantor and Burali-Forti show that there is no universal set and no set containing all ordinal numbers. The semantic paradoxes cannot even be formulated, since they involve notions not expressible within the system.

For the absolute novice a summary will be given here of some of the basic ideas and results used in the text. The reader is urged to skip these explanations now, and, if necessary, to refer to them later on.

A *set* is a collection of objects.† The objects in the collection are called *elements* or *members* of the set, and we shall write "$x \in y$" for the statement that x is a member of y. (Synonymous expressions are "x belongs to y" and "y contains x".) The negation of "$x \in y$" will be written "$x \notin y$".

By "$x \subseteq y$" we mean that every member of x is also a member of y, or, in other words, that x is a *subset* of y (or, synonymously, that x is *included* in y). We shall write "$t = s$" to mean that "t" and "s" denote the same object. As usual, "$t \neq s$" is the negation of "$t = s$". For sets x and y, we assume that $x = y$ if and only if $x \subseteq y$ and $y \subseteq x$; that is, if and only if x and y have the same members. A set x is called a *proper* subset of a set y, written "$x \subset y$", if $x \subseteq y$ but $x \neq y$.‡

The union $x \cup y$ of sets x and y is defined to be the set of all elements which are members of x or y or both. Hence, $x \cup x = x$, $x \cup y = y \cup x$, and $(x \cup y) \cup z = x \cup (y \cup z)$. The *intersection* $x \cap y$ is the set of elements which x and y have in common. It is easy to verify that $x \cap x = x$, $x \cap y = y \cap x$, $x \cap (y \cap z) = (x \cap y) \cap z$, $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$, and $x \cup (y \cap z) = (x \cup y) \cap (x \cup z)$. The *relative complement* $x - y$ is the set of members of x which are not members of y. We also postulate the existence of the *empty set* (or *null set*) 0, i.e., a set which has no members at all. Then, $x \cap 0 = 0$, $x \cup 0 = x$, $x - 0 = x$, $0 - x = 0$, and $x - x = 0$. Two sets x and y are called *disjoint* if $x \cap y = 0$.

Given any objects $b_1, \ldots, b_k$, the set which contains $b_1, \ldots, b_k$ as its only members is denoted $\{b_1, \ldots, b_k\}$. In particular, $\{x, y\}$ is a set having x and y as its only members and, if $x \neq y$, is called the *unordered pair* of x and y. The set $\{x, x\}$ is written $\{x\}$ and is called the *unit set* of x. Notice that $\{x, y\} = \{y, x\}$. On the other hand, by $\langle b_1, \ldots, b_k \rangle$ we mean the *ordered k-tuple* of $b_1, \ldots, b_k$. The basic property of ordered k-tuples is that $\langle b_1, \ldots, b_k \rangle = \langle c_1, \ldots, c_k \rangle$ if and only if $b_1 = c_1, b_2 = c_2, \ldots, b_k = c_k$. Thus, $\langle b_1, b_2 \rangle = \langle b_2, b_1 \rangle$ if and only if $b_1 = b_2$. Ordered 2-tuples are called ordered *pairs*. If X is a set and k is a positive integer, we denote by X^k the set of all ordered k-tuples $\langle b_1, \ldots, b_k \rangle$ of elements $b_1, \ldots, b_k$ of X. We also make the convention that X^1 stands for X. X^k is called the *Cartesian product* of X with itself k times. If Y and Z are sets, then by $Y \times Z$ we denote the set of all ordered pairs $\langle y, z \rangle$ such that $y \in Y$ and $z \in Z$. $Y \times Z$ is called the Cartesian product of Y and Z.

†Which collections of objects form sets will not be specified here. Care will be exercised to avoid using any ideas or procedures which may lead to the paradoxes; all the results can be formalized in the axiomatic set theory of Chapter 4. The term "class" is sometimes used as a synonym for "set", but it will be avoided here because it has a different meaning in Chapter 4. If the property $P(x)$ does determine a set, this set is often denoted $\{x|P(x)\}$.

‡The notation $x \subsetneq y$ is often used instead of $x \subset y$.

An *n-place relation* (or a *relation with n arguments*) on a set X is a subset of X^n, i.e., a set of ordered n-tuples of elements of X. For example, the 3-place relation of betweenness for points on a line is the set of all 3-tuples $\langle x, y, z \rangle$ such that the point x lies between the points y and z. A 2-place relation is called a *binary* relation, e.g., the binary relation of fatherhood on the set of human beings is the set of all ordered pairs $\langle x, y \rangle$ such that x and y are human beings and x is the father of y. A 1-place relation on X is a subset of X, and is called a *property* on X.

Given a binary relation R on a set X, the *domain* of R is defined to be the set of all y such that $\langle y, z \rangle \in R$ for some z; the *range* of R is the set of all z such that $\langle y, z \rangle \in R$ for some y; and the *field* of R is the union of the domain and range of R. The *inverse* relation R^{-1} of R is the set of all ordered pairs $\langle y, z \rangle$ such that $\langle z, y \rangle \in R$. For example, the domain of the relation $<$ on the set ω of non-negative integers† is ω, its range is $\omega - \{0\}$, and the inverse of $<$ is $>$. Notation: Very often xRy is written instead of $\langle x, y \rangle \in R$. Thus, in the example just given, we usually write $x < y$ instead of $\langle x, y \rangle \in <$.

A binary relation R is said to be *reflexive* if xRx for all x in the field of R. R is *symmetric* if xRy implies yRx, and R is *transitive* if xRy and yRz imply xRz. Examples: The relation $\leqslant$ on the set of integers is reflexive and transitive but not symmetric. The relation "having at least one parent in common" on the set of human beings is reflexive and symmetric but not transitive.

A binary relation which is reflexive, symmetric, and transitive is called an *equivalence relation*. Examples of equivalence relations: (1) the *identity relation* I_X on a set X consisting of all pairs $\langle y, y \rangle$, where $y \in X$; (2) the relation of parallelism between lines in a plane; (3) given a fixed positive integer n, the relation $x \equiv y \pmod{n}$ holds when x and y are integers and $x - y$ is divisible by n; (4) the relation between directed line segments in three-dimensional space which holds when and only when they have the same length and the same direction; (5) the congruence relation on the set of triangles in a plane; (6) the similarity relation on the set of triangles in a plane. Given an equivalence relation R on a set X, and given any $y \in X$, define $[y]$ as the set of all z in X such that yRz. Then $[y]$ is called the *R-equivalence class* of y. It is easy to check that $[y] = [z]$ if and only if yRz and that, if $[y] \neq [z]$, then $[y] \cap [z] = 0$, i.e., different R-equivalence classes have no elements in common. Hence, the set X is completely partitioned into the R-equivalence classes. For some of the examples above: (1) the equivalence classes are just the unit sets $\{y\}$, where $y \in X$; (2) the equivalence classes can be considered to be the directions in the plane; (3) there are n equivalence classes, the k^{th} equivalence class ($k = 0, 1, \ldots, n - 1$) being the set of all numbers which leave the remainder k upon division by n; (4) the equivalence classes are the three-dimensional vectors.

†ω will also be referred to as the set of *natural numbers*.

A *function* f is a binary relation such that $\langle x, y\rangle \in f$ and $\langle x, z\rangle \in f$ imply $y = z$. Thus, for any element x of the domain of a function f, there is a unique y such that $\langle x, y\rangle \in f$; this unique element y is denoted $f(x)$. If x is in the domain of f, then $f(x)$ is said to be defined. A function f with domain X and range Y is said to be a *function from X onto Y*. If f is a function from X onto Y, and $Y \subseteq Z$, then f is called a *function from X into Z*. For example, if $f(x) = 2x$ for every integer x, f is a function from the set of integers onto the set of even integers, and f is a function from the set of integers into the set of integers. A function the domain of which consists of n-tuples is said to be a *function of n arguments*. A *(total) function of n arguments on a set X* is a function f whose domain is X^n. We usually write $f(x_1, \ldots, x_n)$ instead of $f(\langle x_1, \ldots, x_n\rangle)$. A *partial* function of n arguments on a set X is a function whose domain is a subset of X^n; e.g. ordinary division is a partial, but not total, function of two arguments on the set of integers (since division by zero is not defined). If f is a function with domain X and range Y, then the *restriction* f_Z of f to a set Z is the function $f \cap (Z \times Y)$. Clearly, $f_Z(u) = v$ if and only if $u \in Z$ and $f(u) = v$. The *image* of the set Z under the function f is the range of f_Z. The *inverse image* of a set W under the function f is the set of all elements u of the domain of f such that $f(u) \in W$. We say that f *maps X* onto (into) Y if X is a subset of the domain of f and the image of X under f is (a subset of) Y. By an *n-place operation* (or *operation with n arguments*) *on a set X* we mean a function from X^n into X. For example, ordinary addition is a binary (i.e., 2-place) operation on the set of natural numbers $\{0, 1, 2, \ldots\}$. But ordinary subtraction is not a binary operation on the set of natural numbers, though it is a binary operation on the set of integers.

Given two functions f and g, the *composition* $f \circ g$ (also sometimes denoted fg) is the function such that $(f \circ g)(x) = f(g(x))$; $(f \circ g)(x)$ is defined if and only if $g(x)$ is defined and $f(g(x))$ is defined. For example, if $g(x) = x^2$ and $f(x) = x + 1$ for every integer x, then $(f \circ g)(x) = x^2 + 1$ and $(g \circ f)(x) = (x + 1)^2$. Also, if $h(x) = -x$ for every real number x and $f(x) = \sqrt{x}$ for every non-negative real number x, then $(f \circ h)(x)$ is defined only for $x \leqslant 0$, and, for such x, $(f \circ h)(x) = \sqrt{-x}$. A function f such that $f(x) = f(y)$ implies $x = y$ is called a 1–1 (*one–one*) function. Examples: (1) The identity relation I_X on a set X is a 1–1 function, since $I_X(y) = y$ for any $y \in X$; (2) the function $g(x) = 2x$, for every integer x, is a 1–1 function; (3) the function $h(x) = x^2$, for every integer x, is not 1–1, since $h(-1) = h(1)$. Notice that a function f is 1–1 if and only if its inverse relation f^{-1} is a function. If the domain and range of a 1–1 function f are X and Y, respectively, then f is said to be a 1–1 (*one–one*) *correspondence between X and Y*; then f^{-1} is a 1–1 correspondence between Y and X, and $(f^{-1} \circ f) = I_X$ and $(f \circ f^{-1}) = I_Y$. If f is a 1–1 correspondence between X and Y, and g is a 1–1 correspondence between Y and Z, then $g \circ f$ is a 1–1 correspondence between X and Z. Sets X and Y are said to be *equinumerous*

(written $X \cong Y$) if and only if there is a 1–1 correspondence between X and Y. Clearly, $X \cong X$; $X \cong Y$ implies $Y \cong X$; and $X \cong Y$ and $Y \cong Z$ imply $X \cong Z$. One can prove (cf. Schröder-Bernstein Theorem, page 194) that if $X \cong Y_1 \subseteq Y$ and $Y \cong X_1 \subseteq X$, then $X \cong Y$. If $X \cong Y$, one sometimes says that X and Y *have the same cardinal number*, and if X is equinumerous with a subset of Y but Y is not equinumerous with a subset of X, one says that the *cardinal number of X is smaller than* the cardinal number of Y.†

A set X is *denumerable* if it is equinumerous with the set of positive integers. A denumerable set is said to have cardinal number $\aleph_0$, and any set equinumerous with the set of all subsets of a denumerable set is said to have the cardinal number $2^{\aleph_0}$ (or to have the *power of the continuum*). A set X is *finite* if it is empty or if it is equinumerous with the set of all positive integers $\{1, 2, \ldots, n\}$ which are less than or equal to some positive integer n. A set which is not finite is said to be *infinite*. A set is *countable* if it is either finite or denumerable. Clearly, any subset of a denumerable set is countable. A *denumerable sequence* is a function s whose domain is the set of positive integers; one usually writes s_n instead of $s(n)$. A *finite sequence* is a function whose domain is $\{1, 2, \ldots, n\}$, for some positive integer n.

Let $P(x, y_1, \ldots, y_k)$ be some relation on the set of non-negative integers. In particular, P may involve only the variable x and thus be a property. If $P(0, y_1, \ldots, y_k)$ holds, and, if, for any n, $P(n, y_1, \ldots, y_k)$ implies $P(n + 1, y_1, \ldots, y_k)$, then $P(x, y_1, \ldots, y_k)$ is true for all non-negative integers x (*Principle of Mathematical Induction*). In applying this principle, one usually proves that, for any n, $P(n, y_1, \ldots, y_k)$ implies $P(n + 1, y_1, \ldots, y_k)$ by assuming $P(n, y_1, \ldots, y_k)$ and then deducing $P(n + 1, y_1, \ldots, y_k)$; in the course of this deduction, $P(n, y_1, \ldots, y_k)$ is called the *inductive hypothesis*. If the relation P actually involves variables $y_1, \ldots, y_k$ other than x, then the proof of "for all x, $P(x)$" is said to proceed by *induction on x*. A similar induction principle holds for the set of integers greater than some fixed integer j. Example: to prove by mathematical induction that the sum of the first n odd integers $1 + 3 + 5 + \ldots + (2n - 1)$ is n^2, first show that $1 = 1^2$ (i.e., $P(1)$), and then, that if $1 + 3 + 5 + \ldots + (2n - 1) = n^2$, then $1 + 3 + 5 + \ldots + (2n - 1) + (2n + 1) = (n + 1)^2$ (i.e., if $P(n)$ then $P(n + 1)$). From the Principle of Mathematical Induction one can prove the *Principle of Complete Induction*: if, for every non-negative integer x the assumption that $P(u, y_1, \ldots, y_k)$ is true for all $u < x$ implies that $P(x, y_1, \ldots, y_k)$ holds, then, for all non-negative integers x, $P(x, y_1, \ldots, y_k)$ is true. (Exercise: show, by complete induction, that every integer greater than 1 is divisible by a prime number.)

†One can attempt to define the cardinal number of a set X as the collection $[X]$ of all sets equinumerous with X. However, in certain systems of set theory, $[X]$ does not exist, whereas in others (cf. page 196), $[X]$ exists but is not a set. For cardinal numbers $[X]$ and $[Y]$, one can define $[X] \leqslant [Y]$ to mean that X is equinumerous with a subset of Y.

A *partial order* is a binary relation R such that R is transitive and, for every x in the field of R, xRx is false. If R is a partial order, then the relation R' which is the union of R and the set of all ordered pairs $\langle x, x\rangle$, where x is in the field of R, we shall call a *reflexive partial order*; in the literature, "partial order" is used for either partial order or reflexive partial order. Notice that (xRy and yRx) is impossible if R is a partial order, while (xRy and yRx) implies $x = y$ if R is a reflexive partial order. A (reflexive) *total order* is a (reflexive) partial order R such that, for any x and y in the field of R, either $x = y$ or xRy or yRx. Examples: (1) the relation $<$ on the set of integers is a total order, while $\leqslant$ is a reflexive total order; (2) the relation $\subset$ on the set of all subsets of the set of positive integers is a partial order, but not a total order, while the relation $\subseteq$ is a reflexive partial order but not a reflexive total order. If C is the field of a relation R, and if B is a subset of C, then an element y of B is called an *R-least element* of B if yRz for every element z of B different from y. A *well-order* (or *well-ordering relation*) is a total order R such that every non-empty subset of the field of R has an R-least element. Examples: (1) the relation $<$ on the set of non-negative integers is a well-order; (2) the relation $<$ on the set of non-negative rational numbers is a total order but not a well-order; (3) the relation $<$ on the set of integers is a total order but not a well-order. Associated with every well-order R having field X there is a corresponding *Complete Induction Principle*: if P is a property such that, for any u in X, whenever all z in X such that zRu have the property P, then u has the property P, then it follows that all members of X have the property P. If the set X is infinite, a proof using this principle is called a proof by *transfinite induction*. One says that a *set X can be well-ordered* if there exists a well-order whose field includes X. An assumption which is useful in modern mathematics but about the validity of which there has been considerable controversy is the *Well-Ordering Principle*: every set can be well-ordered. The Well-Ordering Principle is equivalent (given the usual axioms of set theory) to the *Axiom of Choice* (*Multiplicative Axiom*): given any set X of non-empty pairwise disjoint sets, there is a set Y (called a *choice set*) which contains exactly one element in common with each set in X.

Let B be a non-empty set, f a function from B into B, and g a function from B^2 into B. Let us write x' for $f(x)$, and $x \cap y$ for $g(x, y)$. Then $\langle B, f, g\rangle$ is called a *Boolean algebra* if and only if the following conditions are satisfied:

(i) $x \cap y = y \cap x$ for all x, y in B.
(ii) $(x \cap y) \cap z = x \cap (y \cap z)$ for all x, y, z in B.
(iii) $x \cap y' = z \cap z'$ if and only if $x \cap y = x$ for any x, y, z in B.

We let $x \cup y$ stand for $(x' \cap y')'$; and we write $x \leqslant y$ for $x \cap y = x$. It is easily proved that $z \cap z' = w \cap w'$ for any w, z in B; we denote the value of $z \cap z'$ by 0. (The symbols $\cap$, $\cup$, 0 should not be confused with the corresponding symbols used in set theory.) We let 1 stand for $0'$. Then: $z \cup z' = 1$ for all z in

B; $\leqslant$ is a reflexive partial order on B; and $\langle B, f, \cup \rangle$ is a Boolean algebra. An *ideal* in $\langle B, f, g \rangle$ is a non-empty subset J of B such that: (1) if $x \in J$ and $y \in J$, then $x \cup y \in J$, and (2) if $x \in J$ and $y \in B$, then $x \cap y \in J$. Clearly, $\{0\}$ and B are ideals. An ideal different from B is called a *proper ideal*. A *maximal ideal* is a proper ideal which is included in no other proper ideal. It can be shown that a proper ideal J is maximal if and only if, for any u in B, $u \in J$ or $u' \in J$. From the Well-Ordering Principle (or the Axiom of Choice) it can be proved that every Boolean algebra contains a maximal ideal, or, equivalently, that every proper ideal is included in some maximal ideal. Example: let B be the set of all subsets of a set X; for $Y \in B$, let $Y' = X - Y$, and for Y, Z in B, let $Y \cap Z$ be the ordinary set-theoretic intersection of Y and Z. Then $\langle B, ', \cap \rangle$ is a Boolean algebra. The 0 of B is the empty set 0, and 1 is X. Given an element u in X, let J_u be the set of all subsets of X which do not contain u. Then J_u is a maximal ideal. For a detailed study of Boolean algebras, cf. Sikorski [1960], Halmos [1963], Mendelson [1970].