A COURSE IN MATHEMATICAL LOGIC

by

J. L. BELL

and

M. MACHOVER

A COURSE IN MATHEMATICAL LOGIC

by

J. L. BELL

London School of Economics and Political Science

and

M. MACHOVER

Chelsea College of Science and Technology

1977

NORTH-HOLLAND PUBLISHING COMPANY — AMSTERDAM · NEW YORK · OXFORD

Library of Congress Catalog Card Number: 75–33890
North-Holland ISBN: 0 7204 28440

Published by:
North-Holland Publishing Company
Amsterdam · New York · Oxford

Sole distributors for the U.S.A. and Canada:
Elsevier North-Holland, Inc.
52 Vanderbilt Avenue
New York, N. Y. 10017

Library of Congress Cataloging in Publication Data

Bell. John Lane.
A course in mathematical logic.

Bibliography: p. 576.
Includes index.
1. Logic, Symbolic and mathematical. I. Machover, Moshé, joint author. II. Title.
QA9.B3953 511′.3 75–33890

Printed in Hungary

ACKNOWLEDGEMENTS

In the course of producing this book we have become indebted to many people. To begin with, we would like to put on record our intellectual debt to those logicians and mathematicians whose work we have expounded: in a work of this kind it would not be feasible to attribute each result to its creator, but we hope that the historical references at the end of each chapter will furnish a general (if sketchy) guide to who proved what.

Our students have been very helpful in furnishing comments and criticism; in this connection we would particularly like to thank Michael Bate, Narciso Garcia and Ali al-Nowaihi.

We are also grateful to our colleagues Daniel Leivant, Dirk de Jongh and Brian Rotman for reading sections of the manuscript and making useful suggestions.

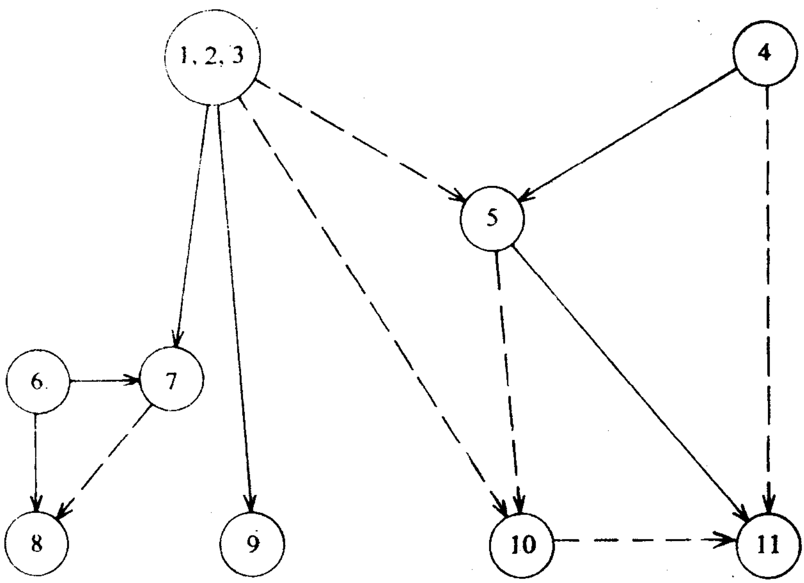
The job of typing the manuscript was, to say the least, an arduous and tricky one and for skilfully carrying out this operation we are indebted to Buffy Fennelly, Diane Roberts, Barbara Silver and Marie-Louise Varichon.

We would also like to thank the printers, and the staff of North-Holland Publishing Company, in particular Einar Fredriksson and Thomas van den Heuvel, for the competent and friendly way they have handled all stages in the production of this book.

Finally, we offer our warmest thanks to our wives Mimi and Ilana for, among many other things, their patience and encouragement throughout the long writing period and for their help in preparing the bibliography.

John Bell
Moshé Machover

INTERDEPENDENCE SCHEME FOR THE CHAPTERS



INTRODUCTION

For the past seven years, the authors have conducted a one-year M.Sc. programme in mathematical logic and foundations of mathematics at London University. The present book developed from our lecture notes for this programme, and the student should therefore be able to work through the text in (roughly) one academic year. The main problem that we faced in constructing the programme was the following. First, we wanted it to be an integrated and balanced account of the most important aspects of logic and foundations. But secondly, since parts of our programme are taken by mathematics and philosophy of science students who for one reason or another do not want to cover all the topics we discuss, we were led to arrange it in such a way that parts could be taken as separate smaller courses. Accordingly, the book itself falls naturally into several units:

1. *Chapters 1–3.* These together constitute an elementary introduction to mathematical logic up to the Gödel–Henkin completeness theorem. We teach this part in a fairly leisurely way (four hours per week for ten weeks, including problem classes), and accordingly the pace of the text here is rather gentle. There is one feature which deserves special mention and that is the use of Smullyan's tableau method. This method serves a dual purpose. First, it is a proof-theoretic instrument that allows us to obtain constructive proofs of various results. In this respect it is equivalent to Gentzen's calculus and to various systems of natural deduction. Secondly, our teaching experience shows that Smullyan's method has the great advantage of being a *practical* tool — after a little practice, it furnishes a quick, efficient and almost computational method of actually detecting the truth or falsehood of formulas. (This efficiency stems in part from the fact that, unlike Gentzen's calculus, it does not require the same formula to be copied again and again.) However, the material

on tableaux has in fact been isolated in separate (starred) sections so that the reader who does not want to use this material can simply ignore it; what remains is a self-contained standard account of first-order logic. A middle course is also possible: a reader wishing to enjoy the *practical* advantages of tableaux but who lacks the time or patience for the somewhat complex constructive proofs of elimination theorems (Ch. 1, §8, and Ch. 2, §§5, 6) can skip the latter because the same results are also obtained in an easier but *non-constructive* way elsewhere (Ch. 1, §9 and Ch. 2, §8). We should like to point out that the somewhat rebarbative complexities of Ch. 2, §5 could have been avoided by using different symbols for free and bound variables (as is often done in texts devoted mainly to proof theory). This, however, would detract slightly from the practical utility of the method and in any case would be contrary to accepted usage in most other branches of logic.

2. *Chapters 4 and 5.* The contents of Chapter 4 are taught for 1 hour a week over 10 weeks, concurrently with the material in Chapters 1–3 (of which it is totally independent). It in fact constitutes a separate short course on Boolean algebras. The material in Chapter 5 — model theory — is taught over the following 10 weeks for 2 hours a week. It depends heavily on Chapter 4 but only slightly on Chapters 1–3 inasmuch as it can be read by anyone having modest acquaintance with the notation and main results of first-order logic.

3. *Chapters 6 and 8.* These two chapters constitute a self-contained course on recursion theory. The material in Chapter 6 is taught for 2 hours a week over 10 weeks, concurrently with Chs. 1–4, of which it is totally independent. There are two points here which call for comment. First, we employ register machines instead of Turing machines, because the former are much closer in spirit to actual digital computers, and are also smoother theoretically. Secondly, this chapter includes a full proof of the Matiyasevich–Robinson–Davis–Putnam (MRDP) theorem that every recursively enumerable relation is diophantine. We believe that — despite the length and tedium of the proof — this result is of such importance that no modern account of recursion theory can afford to omit it. In teaching this part of the chapter, we have found that some of the material in §§13, 14, and the first half of §15 can be omitted in class and given to the student to study at home. As for Chapter 8, it is taught for the following eight weeks at a rate of 2 hours per week. The material here of course depends entirely on Chapter 6, but in this book it appears *after* Chapter 7,

because it is motivated by and illuminates the results contained there. However, no *detailed* knowledge of Chapter 7 is required to understand Chapter 8. In Chapters 6 and 8 we have adopted a somewhat formal approach: in proving that such-and-such a function is recursive, we employ the precise apparatus furnished by the recursion theorem, rather than the intuitive "proof by Church's thesis". We have chosen this course because we believe that the beginning student has not yet developed sufficient experience in the subject to be totally convinced by intuitive proofs which employ Church's thesis.

4. *Chapter 7.* This chapter contains an account of the limitative results about formal mathematical systems. Reliance on the MRDP theorems has enabled us to simplify some of the proofs and obtain somewhat sharper results than usual. The chapter presupposes a good knowledge of first-order logic and some knowledge of recursion theory. However, it can be and is taken by students who have *no* detailed acquaintance with the latter. We have found it feasible to develop all the requisite results from recursion theory — *except* the MRDP theorem — using Church's thesis, the MRDP theorem itself being stated without proof. This approach enables us to teach the material in this chapter intelligibly to students who do not want take a full-fledged course in recursion theory. (The material here is in fact taught concurrently with Chapter 5 for 10 weeks at 2 hours per week.)

5. *Chapter 9.* Here we have an elementary introduction to first-order intuitionistic logic. While neither of the authors claim to be an expert on intuitionism, we nevertheless believe that the constructivist approach to mathematics is of such great importance that some discussion of it is essential. (The material in this chapter is taught concurrently with Chapters 5, 7 and 8 for 10 weeks at 1 hour per week.)

6. *Chapter 10.* This is devoted to an axiomatic investigation of Zermelo–Fraenkel set theory, up to the relative consistency of the axiom of choice and the generalized continuum hypothesis. It requires modest familiarity with first-order logic and with the Löwenheim–Skolem theorem in Chapter 5. (This material is taught over roughly 10 weeks at 2 hours per week at the end of the second term and the beginning of the final term.)

7. *Chapter 11.* This chapter contains an introduction to nonstandard analysis, which is an important method of applying model theory to mathematics. The material here is taught over 10 weeks at 2 hours per week,

during the latter part of the year. Although this chapter presupposes a few results of model theory, these results can be stated concisely without proof for the benefit of those students who wish to study the subject without doing a special course on model theory. In fact it is possible to teach nonstandard analysis to students who have only a slender acquaintance with logic.

As can be seen from the foregoing synopsis, the material in the book can be regarded as forming several relatively independent units. However, the book has been conceived as an organic whole, and provides what is in our view a "balanced diet". We have striven to reveal the interplay between "structural" (i.e. set-theoretical) ideas and "constructive" methods. The latter play a particularly prominent role in mathematical logic, and we have therefore stressed the constructive approach where appropriate but without, we hope, undue fanaticism.

The *problems* constitute an essential part of the book. They are not mere brainteasers, nor should they be too difficult for the student to solve, given the hints that are provided. Many of them contain results which are later employed in proofs of theorems. Accordingly, no unstarred problem should be skipped!

Certain sections and problems are *starred*. This does not necessarily indicate that they are more difficult, but rather that they may be omitted at a first reading. Some problems have been starred because they require more knowledge or skill than is needed for understanding the text in the same section.

Each chapter is divided into sections. When we want to refer to a theorem, problem, definition, etc., within the same chapter, we give the number of the section in which it occurs, followed by its number in that section. Thus, e.g., Def. 2.10.1 is the first numbered statement in §10 of Ch. 2 and within Ch. 2 it is referred to as "Def. 10.1." (or simply "10.1").

We use the convenient abbreviation "iff" for "if and only if". The mark ∎ is used *either* to signify the end of a proof *or*, when it appears immediately after the statement of a result, to indicate that the proof is immediate and is accordingly omitted.

References to the bibliography are given thus: KELLEY [1955]. The overwhelming majority of references to the bibliography are given in a separate section at the end of each chapter.

CONTENTS

CHAPTER 0

PREREQUISITES

In this book we assume that the reader is familiar with the basic facts of naive set theory (including the fundamentals of cardinal and ordinal arithmetic) as presented, e.g., in FRAENKEL [1961], HALMOS [1960] or KURATOWSKI–MOSTOWSKI [1968]. Facts about cardinals and ordinals are used at the end of Ch. 3, occasionally in Chs. 4 and 9, and throughout Chs. 5 and 10. In some places (especially in Chs. 4 and 11), we assume a slender acquaintance with the basic notions of general topology as presented, e.g., in Ch. 1 of BOURBAKI [1961] or the first few chapters of KELLEY [1955].

We distinguish between *classes* and *sets*. Except in Chs. 10 and 11 (where the terms "class" and "set" are assigned a more precise technical meaning), a *class* is understood to be an arbitrary collection of objects, while a *set* is a class which can be a member of another class. (Another distinguishing feature of sets is that only they have cardinalities.)

Given an object x and a class X, we write as usual $x \in X$ for "x is a *member (element, point)* of X and say "*X contains x*" or "x is *in* X". If X contains every member of a class Y, we say "*X includes Y*" and write $Y \subseteq X$. Two classes are regarded as identical if they have the same members.

The set of natural numbers (which contains 0) is denoted by N or ω. Except in Ch. 10, the empty set is denoted by $\emptyset$. If A is a set, the *power set* PA of A is the set of all subsets of A.

Given $n \geqslant 1$ objects $x_1,\ldots,x_n$, we write $\langle x_1,\ldots,x_n\rangle$ for the *ordered n-tuple* of $x_1,\ldots,x_n$. Thus $\langle x,y\rangle$ is the *ordered pair* of x and y. *By convention*, we put $\langle x\rangle = x$ (the ordered singleton of x).

The *Cartesian product* of a finite sequence of classes $A_1,\ldots,A_n$ (with $n \geqslant 1$), denoted by $A_1 \times \ldots \times A_n$, is the collection of all n-tuples $\langle a_1,\ldots,a_n\rangle$ with $a_1 \in A_1,\ldots,a_n \in A_n$. If each A_i is identical with a fixed class A, we write A^n for $A_1 \times \ldots \times A_n$. *By convention*, we set $A^0 = \{\emptyset\}$; thus A^0 has exactly one member, namely $\emptyset$.

For $n \geq 1$, an *n-ary relation* on a class A is a collection of n-tuples of members of A, i.e. a subclass of A^n. A unary relation on A is called a *property*; it is just a subclass of A. The *identity* (or *diagonal*) relation on A is the binary relation

$$\{\langle x,x\rangle : x \in A\}.$$

The *membership* relation on A is the binary relation

$$\{\langle x,y\rangle : x \in A \quad \text{and} \quad y \in A \quad \text{and} \quad x \in y\}.$$

If R is an n-ary relation on A and $B \subseteq A$, the *restriction* of R to B is defined to be the n-ary relation $R \cap B^n$ on B. If R is a binary relation, we often write xRy for $\langle x,y\rangle \in R$.

A *function (map, mapping)* is a class f of ordered pairs such that, whenever $\langle x,y\rangle \in f$ and $\langle x,z\rangle \in f$, we have $y=z$. The *domain* $\mathrm{dom}(f)$ of f is the class

$$\{x : \text{for some } y, \langle x,y\rangle \in f\}$$

and the *range* $\mathrm{ran}(f)$ of f is the class

$$\{y : \text{for some } x, \langle x,y\rangle \in f\}.$$

If f is a function, and $x \in \mathrm{dom}(f)$, then the unique y for which $\langle x,y\rangle \in f$ is denoted (*except in Ch.* 10) by $f(x)$, or sometimes fx, etc., and is called the *value* of f at x. Sometimes we specify a function f in terms of its values; under these conditions we write $x \mapsto f(x)$. (Thus, for example, $x \mapsto x+1$ describes the successor function on N.) If f is a function such that $\mathrm{dom}(f)=A$ and $\mathrm{ran}(f) \subseteq B$, we say that f is a *function from A to (into) B* and write $f : A \to B$. If $f : A \to B$ and $X \subseteq A$, we define the *restriction* $f|X : X \to B$ by

$$(f|X)(x) = f(x) \quad \text{for} \quad x \in X.$$

If $X \subseteq A$ and $Y \subseteq B$, we put

$$f[X] = \{f(x) : x \in X\}, \qquad f^{-1}[Y] = \{x : f(x) \in Y\},$$

and, for $y \in Y$, we put

$$f^{-1}(y) = f^{-1}[\{y\}].$$

A function $f : A \to B$ is *one-one* (an *injection*) if $f(x)=f(y)$ implies $x=y$ for all $x,y \in A$; *onto* (a *surjection*) if $f[A]=B$; and a *one-one correspondence* (a *bijection*) between A and B if both of these conditions hold. A and B are said to be *equipollent* if there is a bijection between A and B. If $f : A \to B$ and $g : B \to C$, the *composition* $g \circ f : A \to C$ is defined by $(g \circ f)(x) = g(f(x))$

for $x \in A$. We sometimes omit the $\circ$ and write simply gf instead of $g \circ f$. Observe that, for each class A, the identity relation on A is a bijection between A and A; for this reason it is also called the *identity map* on A. If $A \subseteq B$, the *natural injection* of A into B is the map $i: A \to B$ defined by $i(x) = x$ for $x \in A$.

If A is a class and I is a set, we write A^I for the collection of all functions from I into A. (Notice that this definition implies $A^\emptyset = \{\emptyset\} = A^0$.) If $\{A_i\colon i \in I\}$ i an indexed family of sets, we write $\prod_{i \in I} A_i$ for the collection of al functions f with domain I such that $f(i) \in A_i$ for all $i \in I$. The *axiom ol choice* asserts that, if each $A_i \neq \emptyset$, then $\prod_{i \in I} A_i \neq \emptyset$.

For any $n \in N$, an *n-ary operation* on a class A is a function from A^n to A. In particular, a 0-ary operation on A is a function from $\{\emptyset\}$ to A, and therefore has a unique value which *we identify with the given* 0-*ary operation*. Thus a 0-ary operation on A is just a member of A. If f is an n-ary operation on A, we write $f(a_1, \ldots, a_n)$ for $f(\langle a_1, \ldots, a_n \rangle)$. A subclass B of A is said to be *closed* or *stable* under f if $f(b_1, \ldots, b_n) \in B$ whenever $b_1, \ldots, b_n \in B$. If B is closed under f, we define the *restriction* $f|B$ of f to B by

$$(f|B)(b_1, \ldots, b_n) = f(b_1, \ldots, b_n) \quad \text{for} \quad b_1, \ldots, b_n \in B.$$

A binary relation R on a class A is called an *equivalence* relation if it satisfies:

(a) xRx for all $x \in A$,

(b) xRy implies yRx for all $x, y \in A$,

(c) xRy and yRz implies xRz for all $x, y, z \in A$.

If R is an equivalence relation on A, for each $x \in A$ the set $x_R = \{y \in A\colon xRy\}$ is called the *R-class* of x. Calling a family $\mathscr{B}$ of subsets of A a *partition* of A if $\bigcup \mathscr{B} = A$ and $X \cap Y = \emptyset$ for any distinct members X, Y of $\mathscr{B}$, we see immediately that, if R is an equivalence relation on A, the family of all R-classes of members of A constitutes a partition of A.

A *partially ordered set* is an ordered pair $\langle A, \leq \rangle$ in which A is a set and $\leq$ is a binary relation on A satisfying:

(a) $x \leq x$ for all $x \in A$,

(b) $x \leq y$ and $y \leq x$ implies $x = y$ for all $x, y \in A$,

(c) $x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in A$.

If $x \leq y$, we say that x is *less than or equal to* y or y is *greater than or equal to* x. We also write "$x < y$" for "$x \leq y$ and $x \neq y$". If $\langle A, \leq \rangle$ is a partially ordered set, $\leq$ is called a *partial ordering* on A. A partially ordered set is said to be *totally* (or *linearly*) *ordered* if in addition

(d) $x \leq y$ or $y \leq x$ for all $x, y \in A$.

If A is any family of sets, the relation $\subseteq$ of set inclusion is a partial ordering on A. We frequently identify a partially ordered set $\langle A,\leqslant\rangle$ with its underlying set A.

If $\langle A,\leqslant\rangle$ is a partially ordered set and $X\subseteq A$, a member $a\in A$ is an *upper (lower) bound* for X, if $x\leqslant a$ $(a\leqslant x)$ for every $x\in X$. An upper (lower) bound a for X in A is called the *supremum (infimum)* of X if a is less than (greater than) every other upper (lower) bound for X in A. If X has a supremum (infimum) in A, we denote it by $\sup(X)$ $(\inf(X))$. Notice that if $\emptyset$ has a supremum (infimum) in A, then $\sup(\emptyset)$ $(\inf(\emptyset))$ is an elemen $a\in A$ such that $a\leqslant x(x\leqslant a)$ for every $x\in A$. That is, if $\sup(\emptyset)$ $(\inf(\emptyset))$ exists in A, then it must be the least (greatest) element of A.

A *chain* in a partially ordered set $\langle A,\leqslant\rangle$ is a subset X of A such that $\leqslant$ totally orders X, i.e. such that $x\leqslant y$ or $y\leqslant x$ for all $x,y\in X$. $\langle A,\leqslant\rangle$ is said to be *inductive* if each chain in A has an upper bound in A. An element $a\in A$ is *maximal* if whenever $x\in A$ and $a\leqslant x$ we have $x=a$. *Zorn's lemma* (which is equivalent to the axiom of choice) asserts that for each element x of an inductive set $\langle A,\leqslant\rangle$ there is a maximal element $a\in A$ such that $x\leqslant a$.

A partially ordered set $\langle A,\leqslant\rangle$ is *well-ordered* if each non-empty subset X of A contains an element x such that $x\leqslant y$ for every $y\in X$. Assuming the axiom of choice, a totally ordered set $\langle A,\leqslant\rangle$ is well-ordered iff A contains no sequence $a_0,a_1,a_2,\ldots$ such that $a_{n+1}<a_n$ for all n.

We conceive of *ordinals* in such a way that each ordinal is the set of all smaller ordinals, and the *finite* ordinals as being identical with the natural numbers. Each well-ordered set is order-isomorphic to a unique ordinal. A *cardinal* is an ordinal which is not equipollent with a smaller ordinal. The *cardinality* of a set X, denoted by $|X|$, is the unique cardinal equipollent with X. (This needs the axiom of choice.) Notice that $|X|=|Y|$ iff X and Y are equipollent. If α and β are cardinals, then α^β denotes the result of *cardinal* exponentation, i.e. the product of α with itself β times. Thus e.g. the cardinality of PA is $2^{|A|}$, for any set A.

A set A is said to be *finite* if it is equipollent with n for some $n\in N$, *denumerable* if it is equipollent with N, and *countable* if it is finite or denumerable.