

**An Introduction to Mathematical Logic and Type Theory:
To Truth Through Proof**

APPLIED LOGIC SERIES

VOLUME 27

Managing Editor

Dov M. Gabbay, *Department of Computer Science, King's College, London, U.K.*

Co-Editor

Jon Barwise†

Editorial Assistant

Jane Spurr, *Department of Computer Science, King's College, London, U.K.*

SCOPE OF THE SERIES

Logic is applied in an increasingly wide variety of disciplines, from the traditional subjects of philosophy and mathematics to the more recent disciplines of cognitive science, computer science, artificial intelligence, and linguistics, leading to new vigor in this ancient subject. Kluwer, through its Applied Logic Series, seeks to provide a home for outstanding books and research monographs in applied logic, and in doing so demonstrates the underlying unity and applicability of logic.

An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof

Second Edition

by

PETER B. ANDREWS

*Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh, Pennsylvania, U.S.A.*



SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-90-481-6079-2 ISBN 978-94-015-9934-4 (eBook)
DOI 10.1007/978-94-015-9934-4

Printed on acid-free paper

All Rights Reserved

© Peter B. Andrews 2002

Originally published by Kluwer Academic Publishers in 2002

Softcover reprint of the hardcover 2nd edition 2002

No part of this work may be reproduced, stored in a retrieval system, or transmitted
in any form or by any means, electronic, mechanical, photocopying, microfilming, recording
or otherwise, without written permission from the Publisher, with the exception

of any material supplied specifically for the purpose of being entered

and executed on a computer system, for exclusive use by the purchaser of the work.

To Cate, Lyle, and Bruce

Contents

Preface to the Second Edition	ix
Preface	xi
0 Introduction	1
1 Propositional Calculus	5
§10. The Language of \mathcal{P}	5
§10A Supplement on Induction	17
§11. The Axiomatic Structure of \mathcal{P}	21
§12. Semantics, Consistency, and Completeness of \mathcal{P}	32
§13. Independence	42
§14. Propositional Connectives	44
§15. Compactness	64
§16. Ground Resolution	67
2 First-Order Logic	73
§20. The Language of \mathcal{F}	73
§21. The Axiomatic Structure of \mathcal{F}	84
§22. Prenex Normal Form	110
§23. Semantics of \mathcal{F}	114
§24. Independence	123
§25. Abstract Consistency and Completeness	125
§25A Supplement: Simplified Completeness Proof	139
§26. Equality	142
3 Provability and Refutability	151
§30. Natural Deduction	151
§31. Gentzen's Theorem	156
§32. Semantic Tableaux	160
§33. Skolemization	166
§34. Refutations of Universal Sentences	171
§35. Herbrand's Theorem	176
§36. Unification	183

4 Further Topics in First-Order Logic	189
§40. Duality	189
§41. Craig's Interpolation Theorem	196
§42. Beth's Definability Theorem	199
5 Type Theory	201
§50. Introduction	201
§51. The Primitive Basis of \mathcal{Q}_0	210
§52. Elementary Logic in \mathcal{Q}_0	215
§53. Equality and Descriptions	232
§54. Semantics of \mathcal{Q}_0	238
§55. Completeness of \mathcal{Q}_0	248
6 Formalized Number Theory	257
§60. Cardinal Numbers and the Axiom of Infinity	257
§61. Peano's Postulates	261
§62. Order	269
§63. Minimization	276
§64. Recursive Functions	281
§65. Primitive Recursive Functions and Relations	289
7 Incompleteness and Undecidability	301
§70. Gödel Numbering	303
§71. Gödel's Incompleteness Theorems	312
§72. Essential Incompleteness	323
§73. Undecidability and Undefinability	332
§74. Epilogue	338
Supplementary Exercises	339
Summary of Theorems	345
Bibliography	371
List of Figures	381
Index	383

Preface to the Second Edition

For the benefit of readers who are already familiar with the first (1986) edition of this book, we briefly explain here how this edition differs from that one.

The basic structure of the book is unchanged, but improvements and enhancements may be found throughout the text. There is more introductory material and motivation. Historical notes and bibliographic references have been added. The discussion of vertical paths and normal forms in §14 is now much more comprehensive. A number of consequences of Gödel's Second Incompleteness Theorem are now discussed in §71.

Many exercises have been added, and at the end of the book there is a new collection of Supplementary Exercises which are not explicitly tied to any particular section. Information about obtaining the ETPS program, which facilitates work on exercises which consist of proving theorems in the object language, will be found near the end of the Preface.

Preface

This book is an introductory text on mathematical logic and type theory. It is aimed primarily at providing an introduction to logic for students of mathematics, computer science, or philosophy who are at the college junior, senior, or introductory graduate level. It can also be used as an introduction to type theory by researchers at more advanced levels.

The first part of the book (Chapters 1 and 2, supplemented by parts of Chapters 3 and 4) is suitable for use as a text in a one-semester introduction to mathematical logic, and the second part (Chapters 5 – 7) for a second semester course which introduces type theory (more specifically, typed λ -calculus) as a language for formalizing mathematics and proves the classical incompleteness and undecidability theorems in this context. Persons who wish to learn about type theory and have had a prior introduction to logic will have no difficulty starting with Chapter 5.

The book is oriented toward persons who wish to study logic as a vehicle for formal reasoning. It may be particularly useful to those who are interested in issues related to the problem of constructing formal proofs as models of informal reasoning or for use in computerized systems which involve automated deduction. Proofs, which are often the chief end products and principal manifestations of mathematical reasoning, constitute highly significant pathways to truth and are a central concern of this book. Our choice of the title *To Truth Through Proof* is motivated by the consideration that while in most realms one needs more than logic to achieve an understanding of what is true, in mathematics the primary and ultimate tool for establishing truth is logic.

Of course, the study of logic involves reasoning about reasoning, and it is not surprising that complex questions arise. To achieve deep understanding and proper perspective, one must study a variety of logical systems as mathematical objects and look at them from a variety of points of view. We are thus led to study the interplay between syntax and semantics, questions of consistency and independence, formal rules of reasoning, various formats for proofs and refutations, ways of representing basic mathematical concepts in a formal system, the notion of computability, and the completeness, incompleteness, and undecidability theorems which illuminate both the power and the limitations of logic.

One of the basic tasks of mathematical logic is the formalization of mathematical reasoning. Both type theory (otherwise known as higher-order logic) and axiomatic set theory can be used as formal languages for this

purpose, and it is really an accident of intellectual history that at present most logicians and mathematicians are more familiar with axiomatic set theory than with type theory. It is hoped that this book will help to remedy this situation.

In logic as in other realms, there is a natural tendency for people to prefer that with which they are most familiar. However, those familiar with both type theory and axiomatic set theory recognize that in some ways the former provides a more natural vehicle than the latter for formalizing what mathematicians actually do. Both logical systems are necessarily more restrictive than naive axiomatic set theory with the unrestricted Comprehension Axiom, since the latter theory is inconsistent. Axiomatic set theory achieves consistency by restricting the Comprehension Axiom and introducing the distinction between sets and classes. Mathematicians often find this distinction unnatural, ignore the technicalities about existence axioms, and leave it to the specialists to show that their reasoning can be justified. Type theory achieves consistency by distinguishing between different types of objects (such as numbers, sets of numbers, collections of sets of numbers, functions from numbers to numbers, sets of such functions, etc.). Mathematicians make such distinctions too, and even use different letters or alphabets to help them distinguish between different types of objects, so the restrictions which type theory introduces are already implicit in mathematical practice. While some formulations of type theory may seem cumbersome, the formulation \mathcal{Q}_0 introduced in Chapter 5 is a very rich and expressive language, with functions (which need not be regarded as sets of ordered pairs) of all types as primitive objects, so most of what mathematicians write can be translated into \mathcal{Q}_0 very directly. \mathcal{Q}_0 is a version of typed λ -calculus, and the availability of λ -notation in \mathcal{Q}_0 enables definitions to be handled very conveniently and eliminates the need for axioms asserting the existence of sets and functions.

One's choice of a formal language will generally depend on what one wishes to do with it. If one is choosing a language for expressing mathematics in computerized systems which deal with mathematics on a theoretical level (such as mathematically sophisticated information retrieval systems, proof checkers, or deductive aids), there are several reasons why type theory may be preferable to set theory. First, the primitive notation of set theory is so economical that this language is only practical to actually use if one expands it by introducing various abbreviations. This is usually done in an informal way (i.e., in the meta-language), but in an automated system it is important that the formal language be the one actually used to express mathematical statements, since this is the language which must be studied rigorously and

manipulated systematically. Thus, the formal language should not only be adequate in principle for expressing mathematical ideas, but it should also be convenient in practice. Indeed, the translation from familiar mathematical notations to the formal language should be as simple and direct as possible. Second, in set theory one can write all sorts of expressions which are not legal in type theory, and which a mathematician would reject as nonsense almost instinctively. A computer program for manipulating mathematical expressions would certainly be very awkward and inefficient if it did not have some way of distinguishing between different types of mathematical objects and rejecting nonsensical expressions. Third, one of the basic operations in automated theorem proving is that of finding appropriate substitutions for variables, and an understanding of the types of various entities allows one to avoid much useless search. Finally, it has been found (see [Andrews *et al.*, 1984] [Andrews *et al.*, 1996]) that unification algorithms for type theory (such as [Huet, 1975] and [Jensen and Pietrzykowski, 1976]) are powerful tools in automating the proofs of simple mathematical theorems.

As noted in [Hanna and Daeche, 1985], [Hanna and Daeche, 1986], and [Gordon, 1986], typed λ -calculus is particularly well suited for specifying and verifying hardware and software. Familiarity with typed λ -calculus also provides fundamental background for the study of denotational semantics for functional programming languages.

While Gödel's Completeness Theorem guarantees that all valid wffs of first-order logic have proofs in first-order logic, extraordinary reductions in the lengths of such proofs can sometimes be achieved by using higher-order logic. (See §54.)

It is easy to discuss the semantics of type theory, and to explain the crucial distinction between standard and nonstandard models which sheds so much light on the mysteries associated with the incompleteness theorems, Skolem's paradox, etc. As will be seen in Chapter 7, in the context of the language Q_0 it is easy to present the incompleteness theorems very elegantly and to show that their significance extends far beyond the realm of formalizing arithmetic.

Obviously, serious students of mathematical logic ought to know about both type theory and axiomatic set theory. Pedagogically, it makes good sense to introduce them to elementary logic (propositional calculus and first-order logic), then type theory, then set theory. In spite of the superficial appearance of simplicity possessed by set theory, its models are more complicated, and from a logical point of view it is a more powerful and complex language. In [Marshall and Chuaqui, 1991] it is argued that the set-theoretical sentences which are preserved under isomorphisms are of fundamental im-

portance, and it is shown that these are the sentences which are equivalent to sentences of type theory.

A brief explanation of the labeling system used in this book may be helpful. The sections in Chapter 1 (for example) are labeled §10, §11, ..., §16. Of course, an alternative labeling would be §1.0, §1.1, ..., §1.6, but the periods are quite redundant once one understands the labeling system. These tags should be regarded as labels, not numbers, although they do have a natural ordering. Similarly, the theorems in §11 (for example) are labeled 1100, 1101, 1102, ..., rather than 1.1.00, 1.1.01, 1.1.02, Exercises are labeled in a similar way, but their labels start with an X. Thus, the exercises for section 11 are labeled X1100, X1101, X1102, Labeling starts with 0 rather than 1 because 0 is the initial ordinal. (Chapter 0 is the Introduction.)

The reader should examine the table of contents in parallel with the discussion of the individual chapters below. Much that is in this book is simply what one would expect in an introductory text on mathematical logic, so the discussion will concentrate on features one might not expect.

Chapter 1 introduces the student to an axiomatic system \mathcal{P} of propositional calculus and later to more general considerations about propositional calculus. In §14 a convenient two-dimensional notation is introduced to represent wffs in negation normal form. This enables one to readily comprehend the structure of such wffs, find disjunctive and conjunctive normal forms by inspection, and check whether the wffs are contradictory by examining the “vertical paths” through them.

Chapter 2 contains what everyone ought to know about first-order logic. Students can learn a great deal about how to construct proofs by doing the exercises at the end of §21. The discussion of prenex normal form in §22 shows students how to pull out the quantifiers of a wff all at once as well as how to bring them out past each connective one step at a time. This discussion is facilitated by the use of the notation \exists' for the ambiguous quantifier.

In §25 Gödel’s Completeness Theorem is proved by an elegant generalization of Henkin’s method due to Smullyan. The notion of consistency in Henkin’s proof is replaced by abstract consistency, and with very little extra effort a proof is obtained for Smullyan’s Unifying Principle, from which readily follow Gödel’s Completeness Theorem in §25, Gentzen’s Cut-Elimination Theorem in §31, the completeness of the method of semantic tableaux in §32, the completeness of a method for refuting universal sentences in §34, and the Craig–Lyndon Interpolation Theorem in §41. In their first encounter with the proof of Gödel’s Completeness Theorem, students may need a less abstract approach than that of Smullyan’s Unifying Prin-

ciple, so as an alternative in §25A an elegant and direct simplified proof of the Completeness Theorem is given. It actually parallels the abstract approach very closely and provides a good introduction to the abstract approach. The Löwenheim–Skolem Theorem and the Compactness Theorem are derived in the usual way from the Completeness Theorem. It is shown how the Compactness Theorem can be applied to extend the Four Color Theorem of graph theory from finite to infinite graphs and to prove the existence of nonstandard models for number theory.

Chapter 3 is concerned with the logical principles underlying various techniques for proving or refuting sentences of first-order logic, and provides an introduction to the fundamental ideas used in various approaches to automated theorem proving.

The chapter starts in §30 with a system of natural deduction which is essentially a summary of the most useful derived rules of inference that were obtained in §21. Methods of proving existential formulas are discussed. A cut-free system of logic and semantic tableaux are then introduced in §31 and §32 as systems in which one can search more systematically for a proof (or refutation). When one constructs a semantic tableau, one often notices that it would be nice to postpone deciding how to instantiate universal quantifiers by just instantiating them with variables for which substitutions could be made later. This is facilitated when one eliminates existential quantifiers to obtain a universal sentence by Skolemization, which is the topic of §33. It is noted that several methods of Skolemization can be used.

§34 introduces a method for refuting universal sentences which avoids the branching on disjunctions of semantic tableaux and uses the cut rule to reduce formulas to the contradictory empty disjunction. This method has many elements in common with the resolution method [Robinson, 1965] of automated theorem-proving. Students are often successful at establishing theorems by this method even though they were unable to find proofs for them in natural deduction format.

The proof methods just discussed all yield highly redundant proof structures in which various subformulas may occur many times. §35 introduces a version of Herbrand’s Theorem which enables one to establish that a wff is contradictory simply by displaying a suitable compound instance of it.

One of the basic processes of theorem-proving in first- or higher-order logic is that of substituting terms for variables (or instantiating quantified variables) in such a way that certain expressions become the same. §36 provides a description of Robinson’s Unification Algorithm for first-order logic and a proof of its correctness. Since substitutions have been defined §10) to be functions of a certain sort which map formulas to formulas, certain

facts about substitutions (such as the associativity of composition) can be used without special justification.

Chapter 4 discusses a few additional topics in first-order logic. In §40 the topic of Duality is introduced with a parable about two scholars who argue about an ancient text on logic because of a fundamental ambiguity in the notations for truth and falsehood. Of course, much of this material could appropriately be discussed immediately after §14.

Chapter 5 introduces a system \mathcal{Q}_0 of typed λ -calculus which is essentially the system introduced by Alonzo Church in [Church, 1940], except that (following [Henkin, 1963b]) equality relations rather than quantifiers and propositional connectives are primitive. Type theory is first introduced in §50 with a discussion of a rather traditional formulation \mathcal{F}^ω of type theory in which propositional connectives and quantifiers are primitive. It is shown that equality can be defined in such a system in two quite natural ways. It is then shown that this system can be simplified to obtain naive axiomatic set theory, which is inconsistent because Russell's paradox can be derived in it. Thus, one needs some device such as type distinctions or restrictions on the Comprehension Axiom to obtain a consistent formalization of naive set theory. The discussion turns to finding as simple, natural, and expressive a formulation of type theory as possible. (The choice of equality as the primitive logical notion is influenced not only by the desire for simplicity but also by the desire for a natural semantics, as discussed at the end of [Andrews, 1972a].) This leads to an exposition of the basic ideas underlying \mathcal{Q}_0 . While these ideas usually come to be regarded as very natural, they often seem novel at this stage, and need to be absorbed thoroughly before one proceeds to the next section.

By the end of §52 the student should be ready to prove theorems in \mathcal{Q}_0 , so notations for some simple but basic mathematical ideas are introduced, and in the exercises the student is asked to give formal proofs of some simple mathematical theorems about sets and functions. An exercise in §53 asks for a proof of a concise type-theoretic formulation of Cantor's Theorem.

Henkin's Completeness Theorem is proved for \mathcal{Q}_0 in §55. Skolem's paradox (which was first discussed in §25) is discussed again in the context of type theory and resolved with the aid of the important distinction between standard and nonstandard models. It is shown that theories which have infinite models must have nonstandard models.

Chapter 6 is intended to make it clear that mathematics really can be formalized within the system \mathcal{Q}_0^∞ which is the result of adding an axiom of infinity to \mathcal{Q}_0 . It is shown how cardinal numbers can be defined, Peano's Postulates can be derived, and recursive functions can be represented very

elegantly in \mathcal{Q}_0^∞ . Proofs of theorems of \mathcal{Q}_0^∞ are presented in sufficient detail so that students should have no difficulty providing formal justifications for each step. Naturally, each teacher will decide how (or whether) to treat these proofs in class. It is good experience for students to present some proofs, or parts of proofs, in class. By the end of this chapter students should have a firm grasp of the connection between abbreviated formal proofs and the informal proofs seen in other mathematics courses, and should therefore be much more confident in dealing with such proofs.

Chapter 7 presents the classical incompleteness, undecidability and undefinability results for \mathcal{Q}_0^∞ . In §70 it is shown that the numerical functions representable in any consistent recursively axiomatized extension of \mathcal{Q}_0^∞ are precisely the recursive functions, and this leads to an argument for Church's Thesis. In §71 Gödel's First and Second Incompleteness Theorems are presented for \mathcal{Q}_0^∞ , along with Löb's Theorem about the sentence which says "I am provable". A number of consequences of Gödel's Second Incompleteness Theorem are discussed. In §72 the Gödel–Rosser Incompleteness Theorem is presented for \mathcal{Q}_0^∞ , and it is shown how this implies that there is no recursively axiomatized extension of \mathcal{Q}_0 whose theorems are precisely the wffs valid in all standard models. §73 is concerned with the unsolvability of the decision problems for \mathcal{Q}_0 and \mathcal{Q}_0^∞ , and the undefinability of truth. §74 is a brief epilogue reflecting on the elusiveness of truth.

The exercises at the end of each section generally provide opportunities for using material from that section. However, students need to learn how to decide for themselves what techniques to use to solve problems. Therefore, at the end of the book there is a collection of Supplementary Exercises which are not explicitly tied to any particular section.

Some exercises (see §21 and §52, for example) involve applying rules of inference to prove theorems of the formal system being discussed. A computer program called ETPS (which is reviewed in [Goldson and Reeves, 1993] and [Goldson *et al.*, 1993]) has been developed to facilitate work on such exercises. The student using ETPS issues commands to apply rules of inference in specified ways, and the computer handles the details of writing the appropriate lines of the proof and checking that the rules can be used in this way. Proofs, and the active lines of the proof, are displayed using the special symbols of logic in proof windows which are automatically updated as the proof is constructed. The program thus allows students to concentrate on the essential logical problems underlying the proofs, and it gives them immediate feedback for both correct and incorrect actions. Experience shows that ETPS enables students to construct rigorous proofs of more difficult theorems than they would otherwise find tractable.

ETPS permits students to work forwards, backwards, or in a combination of these modes, and provides facilities for rearranging proofs, deleting parts of proofs, displaying only those parts of proofs under active consideration, saving incomplete proofs, and printing proofs on paper. A convenient formula editor permits the student to extract needed formulas which occur anywhere in the proof, and build new formulas from them. Teachers who set up ETPS for use by a class can take advantage of its facilities for keeping records of completed exercises and processing information for grades.

Information about obtaining ETPS without cost can be obtained from <http://gtps.math.cmu.edu/tps.html>; alternatively, connect to the web page <http://www.cmu.edu/> for Carnegie Mellon University or <http://www.cs.cmu.edu/afs/cs/project/pal/www/pal.html> for CMU's Pure and Applied Logic Program and from there find the author's web page.

I have used versions of this book for a course in mathematical logic at Carnegie Mellon University for about thirty years. Experience has shown that students are best prepared for the course if they have had at least one rigorous mathematics course or a course in philosophy or computer science that has prepared them to appreciate the importance of understanding the nature of deductive reasoning and the art of proving theorems. I have found that if one covers the material rather thoroughly in class, it is difficult to cover all of Chapters 1 – 4 in the first semester, so I normally cover Chapter 1, most of Chapter 2 (using §25A and omitting §24 and most of §26), and also §30, §33, and §34. Naturally, in a course primarily composed of graduate students, or students with more previous exposure to logic, one could move faster, ask students to read some material outside class, and cover all of Chapters 1 – 4 in one semester. In the second semester I normally cover Chapters 5 – 7 fairly completely.

A few comments about notation may be helpful. “ U is a subset of V ” is written as “ $U \subseteq V$ ”, and “ U is a proper subset of V ” as “ $U \subset V$ ”. The composition of functions f and g is written “ $f \circ g$ ”. “iff” is an abbreviation for “if and only if”. Ends of proofs are marked with the sign ■.

I wish to thank the many people who contributed directly or indirectly to the development of this book. Alonzo Church, John Kemeny, Raymond Smullyan, and Abraham Robinson taught me logic. It will be obvious to all who have read them that this book has been deeply influenced by [Church, 1940] and [Church, 1956]. Numerous students at Carnegie Mellon helped shape the book with their questions, comments, and interests. Special thanks go to Ferna Hartman for her heroic work typing the difficult manuscript.

Chapter 0

Introduction

The ability to reason is one of the marvels of human nature. In many situations it is not easy to think logically, but the benefits of doing so are a familiar fact of human experience. Since at least the time of Aristotle (384–322 B.C.), many questions have been raised by inquisitive and reflective minds about the nature of reasoning, and about ways of improving our use of reason. From such questions and attempts to answer them has arisen the intellectual discipline known as logic.

Let us put the study of logic into perspective. Humans are born with hands, but the development of tools has vastly expanded what humans can do. Humans are born with eyes, but the study of optics and other relevant disciplines has made it possible to develop sophisticated telescopes and microscopes which vastly expand what we can see, leading to amazing advances in fields such as astronomy and biology. We are endowed with the ability to reason, and the study of logic can serve the following purposes:

- (1) enhance our ability to reason correctly, deeply, and with well founded confidence;
- (2) prepare us to use, and perhaps help develop, the automated reasoning tools which are being developed and which will have revolutionary impact on many intellectual activities;
- (3) prepare us to study the many deep and fascinating questions in mathematics, philosophy, and computer science which have their roots in logic.

Note that as the development of automated reasoning tools makes it increasingly practical to use very rigorous reasoning in various disciplines, there will be increasing demand that such standards of rigor be adhered to. For example, the day will come when computer hardware and software

which has simply been tested on many cases will be regarded as unacceptably unreliable. Only hardware and software which has been proved to meet its specifications will be regarded as acceptable.

Intellectual disciplines, such as scientific or philosophical theories, usually evolve over many years, often many centuries. Starting with isolated statements expressing various facts, they gradually become more organized and systematic as people struggle to find order in the phenomena under consideration. Gradually, general principles emerge which enable one to understand many particular facts as manifestations of these general principles. If the process progresses sufficiently far, one may obtain a systematic presentation of these general principles in which all of them can be derived from a few very basic and well specified principles called the *postulates* of the theory that has developed. If one goes a little farther, one specifies not only the postulates, but also the exact language in which the theory is to be expressed and the rules of logical reasoning which are to be used. One thus obtains what logicians call a *formal theory*, or *logistic system*. It can be seen that formalization of a theory is the ultimate act of systematization in a long chain of developments. It removes some of the last potential sources of confusion and vagueness. When one formalizes a theory, one may clarify what assumptions are tacitly made by those who habitually use the theory. (The Axiom of Choice is a striking example of an assumption which was used tacitly in mathematics for a number of years before it was explicitly recognized.¹)

The development of automated reasoning tools makes it likely that formalization of theories will eventually be regarded as a practical necessity, since such formalization will facilitate the development of comprehensive and sophisticated information systems which incorporate scientific and technical theories, and which can derive relevant consequences of these theories. The ability to derive logical consequences from fundamental principles provides an enormously economical way of storing information. Indeed, it permits one to obtain information which was never explicitly put into the system, and obtain it in a variety of forms.

Of course, attempts to understand the nature of reasoning and to build sophisticated information systems which can draw logical conclusions may be regarded as part of an endeavor to fashion more powerful intellectual tools for coping with the increasingly complex problems which confront mankind.

It is sometimes felt that in order to make certain disciplines (such as social sciences) more scientific, one should make them more "mathematical".

¹See [Moore, 1982, Chapter 1].

To the unsophisticated observer of mathematics, it may appear that the use of numbers is the critical characteristic of disciplines which use mathematics, and that one can thus make a discipline more scientific by involving numbers in it more deeply. However, the really fundamental feature of mathematics is not numbers but logical rigor. While arguments involving numbers are often rigorous, it will be seen that a standard of rigor can be defined and maintained whether numbers are involved or not.

Our study of logic will be concerned with only certain aspects of the reasoning process. Our focus will not be on the psychology or physiology of human reasoning, but on reasoning as an abstract process which can be used to derive new information from known information.

A striking feature of the reasoning process is the “magic leap” which a reasoner makes from what is known to conclusions which “follow logically” from what is known. As we shall see, this magic leap can be analyzed and studied in a mathematically precise way.

How can something as abstract and apparently nebulous as thoughts be analyzed with mathematical precision? It turns out that there is a correspondence between our thoughts and the words we use to express them, and between our logical processes and manipulations of strings of words. Thus, we can analyze logical inferences in terms of operations on symbolic expressions, and these can be defined precisely and treated as objects of mathematical study.

Indeed, it has been found that valid arguments can be justified by showing that they have correct logical forms, even if one does not understand the meanings of many of the words in the arguments. It is important to learn to focus on the form rather than the content of an argument when one is analyzing its correctness, because in complicated situations one can easily be led into error by relying on one’s intuition about what is true instead of using careful logical deduction.

Chapter 1

Propositional Calculus

§10. The Language of \mathcal{P}

As discussed in the Introduction, we can study the nature of reasoning by studying logistic systems in which reasoning is represented in a precise way. We shall commence our study with a rather simple logistic system called \mathcal{P} , which is one formulation of propositional calculus. Many of the definitions and concepts which we shall introduce while studying \mathcal{P} will also be applicable to richer logical systems. In addition, \mathcal{P} will occur as a sub-system of some of these richer systems. Once we have discussed \mathcal{P} , which is an example of a logistic system, we will be in a good position to give a general explanation of what a logistic system is.

In order to make clear what aspects of reasoning will be represented in \mathcal{P} , we start by considering three examples of logical inferences.

EXAMPLE 1: Jack is asked whether he will go to the picnic on Saturday. He doesn't like to go to picnics in the rain, and he remembers that if it does not rain, he will be playing in a tennis match which conflicts with the picnic. Hence he replies,

- (A₁) "If it rains, I will not go to the picnic."
- (B₁) "If it does not rain, I will not go to the picnic."
- (C₁) "Therefore, I will not go to the picnic."

EXAMPLE 2: We are given that w , z , and y are sets such that $z - y = \emptyset$ and $w \cap \tilde{z} \subseteq y$ (where \tilde{z} is the complement of the set z), and we wish to show

that $w \subseteq y$. So we assume that $z - y = \emptyset$ and $w \cap \tilde{z} \subseteq y$ and $x \in w$, and try to prove $x \in y$. Using the information that is given, it is not hard to show that

(A₂) If $x \in z$, then $x \in y$.

(B₂) If $x \notin z$, then $x \in y$.

(C₂) Therefore, $x \in y$.

EXAMPLE 3: Let us use $\star\star$ as a name for the sentence “The first sentence in this chapter which contains an occurrence of the symbol @ is not true.” One may argue that:

(A₃) If $\star\star$ is true, then $\star\star$ is not true.

(B₃) If $\star\star$ is not true, then $\star\star$ is not true.

(C₃) Therefore, $\star\star$ is not true.

The reader need not be concerned with all the details of the arguments above, or even with whether the statements (A_i) and (B_i) were correctly inferred in each argument. The important point is that (C_i) is indeed a logical consequence of (A_i) and (B_i) in each argument.

Note that the main inference in each argument has the following form:

(A) If p , then q .

(B) If not p , then q .

(C) Therefore, q .

Obviously, the key words involved in this inference are “if”, “then”, and “not”. These are connectives, words which can be used to construct more complex statements from simpler statements. Statements are sometimes called propositions (though some prefer to say that statements express propositions), and certain connectives which are important for logical purposes are represented in logistic systems as *propositional connectives*. A propositional calculus is a logistic system which formalizes the logical use of propositional connectives.

It is customary to use special symbols instead of English words for propositional connectives. This is very convenient, since the connectives are used very often and the symbols are quicker to write than the words (once one

becomes accustomed to them). Also, the symbols allow one to exhibit very clearly the logical structure of a proposition which might be expressed in a variety of ways in English. Logistic systems which use symbols in this way are also referred to as systems of symbolic logic.

We shall express “If p , then q ” symbolically as “ $p \supset q$ ”, and “not p ” as “ $\sim p$ ”. In this notation, the inference above can be expressed as follows:

$$(A') \ p \supset q$$

$$(B') \ \sim p \supset q$$

$$(C') \ q$$

(The word “therefore” indicates that a conclusion is being reached, but has no logical content itself, so we have omitted it from (C') .)

Before discussing the details of the logistic system \mathcal{P} , let us make a brief survey of the most important propositional connectives.

Negation (\sim)

We write the negation of the statement p as $\sim p$. In English this is generally expressed by inserting the word “not” into the sentence at an appropriate point. If w is the statement “everyone gets wet”, $\sim w$ can be expressed by the statement “not everyone gets wet”, but if r is the statement “it is raining”, $\sim r$ is expressed by “it is not raining”.

If p is true, then the statement $\sim p$ is false, while if p is a false statement, then $\sim p$ is a true statement. This is summarized in Figure 1.1. which we call the truth table for \sim . We write T for True, and F for False.

p	$\sim p$
T	F
F	T

Figure 1.1: Truth table for \sim

Conjunction (\wedge)

The conjunction of statements p and q is written as $[p \wedge q]$, and is generally expressed in English as “ p and q ” or “ p , but q ”. Note that the statements “It is raining, and Jack is going to the picnic” and “It is raining, but Jack is going to the picnic” are *logically equivalent*, though they differ in emphasis. They both mean that the statement “It is raining” is true and

p	q	$p \wedge q$	$p \vee q$	$p \supset q$	$p \equiv q$	$p \not\equiv q$
T	T	T	T	T	T	F
T	F	F	T	F	F	T
F	T	F	T	T	F	T
F	F	F	F	T	T	F

Figure 1.2: Truth tables for \wedge , \vee , \supset , \equiv , $\not\equiv$

the statement "Jack is going to the picnic" is true. The statement $p \wedge q$ is true precisely when p and q are both true.

In Figure 1.2 we display the truth table for \wedge and the other connectives which we will discuss below. Each horizontal line of the truth table corresponds to one of the logical possibilities with regard to the truth or falsity of the statements p and q . For example, in line 1 (below the bar) of the truth table, p is true and q is true and $[p \wedge q]$ is true, while in line 2, p is true and q is false and $[p \wedge q]$ is false.

Disjunction (\vee)

The disjunction of p and q is written as $[p \vee q]$, and generally expressed in English as " p or q ". As can be seen from Figure 1.2, we regard $[p \vee q]$ as true if either or both of the statements p and q are true. Thus, we use \vee for *inclusive* disjunction; we include the case where p and q are both true as one of the cases in which $[p \vee q]$ is true. For mathematical purposes, this is the generally accepted usage of the word "or". For example, 24 is regarded as a member of the set of integers which are multiples of 2 or multiples of 3. However, in ordinary usage the word "or" is sometimes used in the *exclusive sense*, in which " p or q " means p is true or q is true, but not both. For example, if Jack says, "I will go to the picnic or I will go to the tennis match", he may mean that he will go to one or the other, but not both. We will discuss exclusive disjunction below.

Implication (\supset)

$[p \supset q]$ means that the statement p implies the statement q , i.e., q is true whenever p is true. This can be translated into English in a variety of ways, including "if p , then q "; " q , if p "; " p only if q "; " p is sufficient for q "; " q is necessary for p " (i.e., "it is necessary for q to be true in order for p to be true").

We intend \supset to represent *material implication*, so that the truth of $[p \supset q]$ depends only on the truth or falsity of the statements p and q .

Other notions of implication, in which the truth of “ p implies q ” may depend on such factors as the relevance of p to q , may be useful for certain purposes, but are much more complex than material implication, and seem not to be necessary for the formalization of mathematics, so we shall not consider them further here.

Note from Figure 1.2 that $[p \supset q]$ is regarded as true when p is false, without regard for q . The appropriateness of this way of defining material implication may be illuminated by supposing that p means “it rains”, and q means “the street gets wet”, so $[p \supset q]$ means “if it rains, then the street gets wet”. If the street is protected from the rain in some way, so that on a certain day it rains, but the street stays dry, then on that day the statement “if it rains, then the street gets wet” is clearly false. However, if the street is not so protected, we would expect the statement “if it rains, then the street gets wet” to be true every day, whether or not it rains that day. In particular, if on a certain day it does not rain but the street gets wet because a pail of water is spilled on it, the statement “if it rains, then the street gets wet” is still true. Similarly, the statement is true on a day when it does not rain and the street remains dry. Thus the only case in which $[p \supset q]$ is false is that in which p is true and q is false.

Equivalence (\equiv)

$[p \equiv q]$ means that the statements p and q are equivalent with respect to truth; that is, they are both true or both false.

Figure 1.3 displays truth tables for $[[p \supset q] \wedge [q \supset p]]$ and $[p \equiv q]$. In each line of the truth table, the truth of each component statement is displayed under the main connective of that statement. It is easy to see that in every possible situation, $[[p \supset q] \wedge [q \supset p]]$ and $[p \equiv q]$ are both true or both false; that is, they are equivalent. Thus, p is equivalent to q means that p implies q and that q implies p . This leads to several alternative ways of expressing the assertion that p is equivalent to q which are encountered very frequently in mathematics.

p	q	$[[p \supset q] \wedge [q \supset p]]$	$[p \equiv q]$
T	T	T	T
T	F	F	F
F	T	F	F
F	F	T	T

Figure 1.3: Truth tables for $[[p \supset q] \wedge [q \supset p]]$ and $[p \equiv q]$

p	q	$\sim [p \equiv q]$	$p \not\equiv q$
T	T	F	T
T	F	T	F
F	T	T	F
F	F	F	T

Figure 1.4: Truth table for $\not\equiv$

If p is equivalent to q , then q implies p , so we may say “ p , if q ”; also p implies q , so we may say “ p only if q ”; combining these, we say “ p , if q , and p only if q ”, which is shortened to “ p if and only if q ”. (For example, in discussing Figure 1.3 we could have said that in each line of the table, $[[p \supset q] \wedge [q \supset p]]$ is true if and only if $[p \equiv q]$ is true.) In mathematical writing it is common to shorten “ p if and only if q ” even further to “ p iff q ”.

Similarly, if p is equivalent to q , then q implies p , so we may say “ p is necessary for q ”; also p implies q , so we may say “ p is sufficient for q ”; combining these, we say “ p is necessary and sufficient for q ”.

Non-equivalence, exclusive disjunction ($\not\equiv$)

As can be seen from Figures 1.4 and 1.2, $\sim [p \equiv q]$ expresses the *exclusive* disjunction of p and q ; it is true precisely when one of the statements p and q is true, while the other is false, and excludes the case where they are both true. $[p \not\equiv q]$ is often expressed in English by saying “ p unless q ”. For example, if Jack says “I will go to the picnic unless it rains”, he probably means that he will go to the picnic and it will not rain, or it will rain and he will not go to the picnic.

Now that we have some familiarity with propositional connectives, we turn our attention to the details of defining the logistic system \mathcal{P} in which those aspects of reasoning which involve propositional connectives can be formally represented, clarified, and studied rigorously. We shall make \mathcal{P} as simple as possible while still being adequate for our purposes.

Certain propositional connectives can be defined in terms of others. Indeed, we saw in Figure 1.3 how \equiv could be defined in terms of \wedge and \supset . We shall see that all propositional connectives can be defined in terms of \sim and \vee , and for the sake of simplicity these will be the only connectives in \mathcal{P} .

Since the exact nature of the statements to which we will apply the propositional connectives need not concern us for the present, we shall introduce variables, called *propositional variables*, to play the role of statements.

This is analogous to using variables like x and y to play the role of arbitrary numbers when discussing laws of arithmetic such as $x + y = y + x$.

The first step in defining the system \mathcal{P} is to define precisely what sequences of symbols may be used as the language of \mathcal{P} . We start by describing the symbols.

The *primitive symbols* of \mathcal{P} are the following:

- (a) Improper symbols: [] ~ ∨
- (b) Proper symbols: denumerably many propositional variables:
 $p q r s p_1 q_1 r_1 s_1 p_2 q_2 \dots$

A *formula* is a finite sequence of primitive symbols. If \mathbf{X} and \mathbf{Y} are formulas, we shall denote by \mathbf{XY} the formula which can be constructed by concatenating the sequence \mathbf{Y} to the end of the sequence \mathbf{X} . Thus, if \mathbf{X} is $\sim [\sim p$ and \mathbf{Y} is q (the sequence of length one whose sole term is q), \mathbf{XY} is $\sim [\sim pq$.

It will sometimes be important to distinguish carefully between the language \mathcal{P} , which is the object of our study, and the language (English supplemented by mathematical terminology and symbolism) in which we discuss \mathcal{P} . The former we call the *object language*, and the latter the *meta-language*. (For example, if a class which is conducted in French is studying Chinese, then the object language is Chinese, and the meta-language is French.) We shall prove theorems in both the object language and the meta-language. Indeed, in the meta-language we shall prove theorems about the theorems of the object language. For emphasis, we may call a theorem of the meta-language a *metatheorem*.

We shall use formulas of the object language as names for themselves in the meta-language. (Indeed, this has already been done above.) We shall sometimes find it convenient to have symbols which stand for arbitrary propositional variables, just as in mathematics one sometimes uses a letter such as x to refer to an arbitrary real number. We shall henceforth use p , q , r and s , with or without subscripts, to stand for arbitrary propositional variables. Technically, they are variables of our meta-language. That portion of the meta-language which is used to describe the syntax of \mathcal{P} is called the *syntax language* for \mathcal{P} , and p , q , etc., are called *syntactical variables* for the propositional variables of \mathcal{P} .

From among the formulas we shall single out for special attention the *well-formed formulas* (*wffs*), which will be the “meaningful expressions” of \mathcal{P} . We shall define the set of wffs inductively, so that a formula is a wff iff its being so is a consequence of the following *formation rules*:

- (1) A propositional variable standing alone is a wff.
- (2) If \mathbf{A} is a wff, then $\sim \mathbf{A}$ is a wff.
- (3) If \mathbf{A} and \mathbf{B} are wffs, then $[\mathbf{A} \vee \mathbf{B}]$ is a wff.

DEFINITION. The *set of wffs* is the intersection of all sets \mathcal{S} of formulas such that:

- (i) $\mathbf{p} \in \mathcal{S}$ for each propositional variable \mathbf{p} .
- (ii) For each formula \mathbf{A} , if $\mathbf{A} \in \mathcal{S}$, then $\sim \mathbf{A} \in \mathcal{S}$.
- (iii) For all formulas \mathbf{A} and \mathbf{B} , if $\mathbf{A} \in \mathcal{S}$ and $\mathbf{B} \in \mathcal{S}$, then $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{S}$.

A *wff* is a member of the set of wffs.

This formal definition of wffs is mathematically precise, and provides a firm foundation for rigorous proofs of metatheorems about wffs. The definition says that a formula is a wff iff it is in *every* set \mathcal{S} which satisfies conditions (i), (ii), and (iii). Clearly, the set of wffs is the “smallest” set \mathcal{S} satisfying conditions (i), (ii), and (iii) if we use “smaller than” in the sense that \mathcal{S}_1 is smaller than \mathcal{S}_2 iff $\mathcal{S}_1 \subseteq \mathcal{S}_2$.

Using this definition, it is easy to prove (Exercise X1000) the formation rules which were stated above.

EXAMPLE: Let \mathcal{S} be any set satisfying conditions (i), (ii), and (iii). Then $p \in \mathcal{S}$ by (i), and $\sim p \in \mathcal{S}$ by (ii), and $[p \vee \sim p] \in \mathcal{S}$ by (iii). Thus $[p \vee \sim p]$ is in every set satisfying the three conditions, so it is a wff.

EXAMPLE: Let \mathcal{S}_1 be the set of all formulas in which propositional variables occur, and let \mathcal{S}_2 be the set of all formulas in which $[$ and $]$ occur the same number of times. It is easy to see that \mathcal{S}_1 and \mathcal{S}_2 both satisfy conditions (i), (ii), and (iii), so they both contain all wffs. The formula $[[\vee] \vee]$ is not in \mathcal{S}_1 (though it is in \mathcal{S}_2), so it is not a wff. The formula $\sim p [\sim q]$ is also not in *all* sets \mathcal{S} satisfying (i), (ii), and (iii) (as seen in Exercise X1004), although it is in both \mathcal{S}_1 and \mathcal{S}_2 , so it is not a wff either.

We shall use **A**, **B**, **C**, **D**, **E**, etc., as syntactical variables for wffs.

As an immediate consequence of the definition of the set of wffs, we have the following useful induction rule:

1000 Principle of Induction on the Construction of a Wff. Let \mathcal{R} be a property of formulas, and let $\mathcal{R}(\mathbf{A})$ mean that \mathbf{A} has property \mathcal{R} . Suppose

- (1) $\mathcal{R}(\mathbf{p})$ for each propositional variable \mathbf{p} .
- (2) Whenever $\mathcal{R}(\mathbf{A})$, then $\mathcal{R}(\sim \mathbf{A})$.
- (3) Whenever $\mathcal{R}(\mathbf{A})$ and $\mathcal{R}(\mathbf{B})$, then $\mathcal{R}([\mathbf{A} \vee \mathbf{B}])$.

Then every wff has property \mathcal{R} .

Proof: Let \mathcal{S} be the set of wffs which have property \mathcal{R} . It is easily seen that \mathcal{S} has properties (i), (ii), and (iii) in the definition of the set of wffs, so every wff is in \mathcal{S} , and therefore has property \mathcal{R} . ■

Induction is used very frequently in logic, mathematics, and computer science, and our inductive definition of the set of wffs, with the associated Principle of Induction embodied in Theorem 1000, provides a typical example which is worth reflecting on. We have a certain set \mathcal{W} (the set of wffs) which we wish to define. \mathcal{W} must contain certain *initial objects* (the propositional variables) and be closed¹ under certain *operations* (negation and disjunction). Moreover, \mathcal{W} is to contain only those objects that can be obtained from the initial objects by applying the operations an arbitrary number of times. We therefore define \mathcal{W} to be the intersection of all sets which contain the initial objects and are closed under the operations. When \mathcal{W} is defined in this way, it contains the initial objects and is closed under the operations. Moreover, there is always a Principle of Induction on the Construction of the Objects in \mathcal{W} analogous to Theorem 1000: any property which is possessed by all the initial objects and preserved by the operations is possessed by all the members of \mathcal{W} .

We will see many examples of this method in this book. We will use it to define the wffs of a variety of logistic systems, and to define the primitive recursive functions and the recursive functions. In Chapter 6, where we show how to define various mathematical concepts in our object language, we will use it to define the set of natural numbers (with 0 as the sole initial object, the successor function as the sole operation, and the familiar Principle of Mathematical Induction as the induction principle), the order relation on the set of natural numbers, and a recursion operator.

Once one knows some basic facts about wffs, the Principle of Induction on the Construction of a Wff can be given an alternative justification in terms of numerical induction, and applications of it can be replaced by arguments involving numerical induction. This is a somewhat clumsier approach, but

¹We say that a set \mathcal{W} is *closed* under an operation if the result of applying the operation to members of \mathcal{W} always produces members of \mathcal{W} . For example, the disjunction of wffs is a wff, so the set of wffs is closed under disjunction.

p	q	$\sim p$	$\vee q$	$p \supset q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Figure 1.5: $[\sim p \vee q]$ and $[p \supset q]$

p	q	\sim	$[\sim p]$	\vee	$\sim q]$	$p \wedge q$
T	T	T	F	F	F	T
T	F	F	F	T	T	F
F	T	F	T	T	F	F
F	F	F	T	T	T	F

Figure 1.6: $\sim [\sim p \vee \sim q]$ and $[p \wedge q]$

it is widely used and may be comforting to those who prefer to do induction on numbers rather than wffs, so we show how it can be done in Supplement §10A.

A wff **B** which occurs as a (consecutive) part of a wff **A** will be known as a *well-formed (wf) part* of **A**. As a special case, **A** is a wf part of itself.

We next wish to introduce several abbreviations which will allow us to readily express in \mathcal{P} ideas which are most naturally expressed with the aid of propositional connectives which are not primitive connectives of \mathcal{P} . Note from Figure 1.5 that $[\sim p \vee q]$ and $[p \supset q]$ have the same truth value in every possible situation. Similarly, Figure 1.6 shows that $\sim [\sim p \vee \sim q]$ and $[p \wedge q]$ have the same truth value in every possible situation. Also recalling Figure 1.3, we are thus led to make the following definitions:

$[A \supset B]$ stands for $[\sim A \vee B]$.

$[A \wedge B]$ stands for $\sim [\sim A \vee \sim B]$.

$[A \equiv B]$ stands for $[[A \supset B] \wedge [B \supset A]]$.

We say that **A** is the *scope* of the exhibited occurrence of \sim in the wff $\sim A$, and that **A** and **B** are the *left* and *right scopes*, respectively, of the exhibited occurrence of \vee in the wff $[A \vee B]$. The scopes of the other connectives are defined similarly. \sim always has the smallest possible scope, so the scope of the exhibited occurrence of \sim in a wff of the form $[\sim p \vee B]$ is **p**.

For the sake of convenience, we introduce the following conventions which permit us to omit brackets when writing wffs and abbreviations for wffs:

- (a) The outermost brackets of a wff may be omitted.
- (b) A square dot (\blacksquare) stands for a left bracket, whose mate is as far to the right as is possible without altering the pairing of left and right brackets already present.²
- (c) In restoring brackets associated with several occurrences of the same connective, use the convention of association to the left, so that the connective on the left has the smaller scope. Thus,
 $A \supset B \supset C$ stands for $[A \supset B] \supset C$ and
 $A \supset B \supset C \supset D$ stands for $[(A \supset B) \supset C] \supset D$.
- (d) When the relative scopes of several connectives of different kinds must be determined, \sim is to be given the smallest possible scope, then \wedge the next smallest possible scope except for \sim , then \vee , then \supset , then \equiv .

In summary, one may proceed as follows to restore brackets to an abbreviated wff, always taking care not to alter the pairing of left and right brackets already present:

- (i) First restore brackets associated with the square dots (\blacksquare) in a wff; place a left bracket in place of the dot, and its mate as far to the right as possible.
- (ii) Secondly, starting from the left, restore the brackets associated with each occurrence of \wedge so as to give these connectives the smallest possible scope without enlarging the scope of any occurrence of \sim . (Thus $\sim p \wedge q$ stands for $[\sim p \wedge q]$ rather than $\sim [p \wedge q]$. Note that by proceeding from left to right one automatically implements the convention of association to the left.)
- (iii)-(v) Similarly restore all brackets associated with occurrences of \vee , then \supset , then \equiv .

Note that the scope of \sim in the wff $[\sim p \vee q \vee r]$ is p , while the scope of \sim in the wff $[\sim [p \vee q]] \vee r$ is $[p \vee q]$, and the scope of \sim in the wff $[\sim \blacksquare p \vee q \vee r]$ is $[p \vee q \vee r]$.

EXAMPLES:

- (1) $p \vee q \supset \sim r \wedge s \supset r$ stands for $[[p \vee q] \supset [\sim r \wedge s]] \supset r$.
- (2) $p \wedge [\sim r \vee \blacksquare q \supset \blacksquare r \vee s] \vee p$ stands for $[[p \wedge [\sim r \vee [q \supset [r \vee s]]]] \vee p]$.

²The use of dots to replace brackets was introduced by Peano and was used by Whitehead and Russell in [Whitehead and Russell, 1913]. The different dot convention we use here was introduced by Alonzo Church and used in [Church, 1940] and [Church, 1956].

- (3) $\sim r \wedge \sim p \vee \neg q \wedge r \vee [\sim s \supset \neg p_1 \wedge q \supset p] \equiv s \vee p \supset s$ stands for
 $[(\sim r \wedge \sim p) \vee ((q \wedge r) \vee (\sim s \supset ([p_1 \wedge q] \supset p)))] \equiv [(s \vee p) \supset s]]$.

Sometimes it is convenient to write certain brackets as parentheses to make it easier to see how brackets are paired. When this is done, brackets which are written as parentheses should always have their mates treated in the same way. For example, we may regard

$$[(p \equiv q) \equiv r] \wedge \sim [p \equiv (q \equiv r)]$$

as an abbreviation for

$$[[p \equiv q] \equiv r] \wedge \sim [p \equiv [q \equiv r]].$$

DEFINITION. A function θ from the set of formulas into the set of formulas is a *substitution* iff

- (1) $\theta\mathbf{X}$ is the empty formula iff \mathbf{X} is the empty formula.
- (2) for all formulas \mathbf{X} and \mathbf{Y} , $\theta(\mathbf{XY}) = (\theta\mathbf{X})(\theta\mathbf{Y})$; i.e., θ applied to a formula \mathbf{XY} which is the concatenation of \mathbf{X} and \mathbf{Y} is the result of concatenating $\theta\mathbf{X}$ and $\theta\mathbf{Y}$.

Clearly a substitution is determined by its behavior on unit formulas; that is, if θ_1 and θ_2 are substitutions such that $\theta_1\mathbf{x} = \theta_2\mathbf{x}$ for every primitive symbol \mathbf{x} , then $\theta_1 = \theta_2$. Note that $\theta\mathbf{x}$ is the formula which we substitute for \mathbf{x} . A substitution θ is *finite* iff the set of primitive symbols \mathbf{x} such that $\theta\mathbf{x} \neq \mathbf{x}$ is finite. For the most part we shall be concerned with *substitutions for variables*, i.e., substitutions θ for which the only primitive symbols \mathbf{x} such that $\theta\mathbf{x} \neq \mathbf{x}$ are variables.

NOTATION. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be distinct primitive symbols and let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be formulas. $S_{\mathbf{Y}_1 \dots \mathbf{Y}_n}^{\mathbf{x}_1 \dots \mathbf{x}_n}$ is that (finite) substitution θ such that $\theta\mathbf{x}_i = \mathbf{Y}_i$ for $1 \leq i \leq n$ and $\theta\mathbf{y} = \mathbf{y}$ for any primitive symbol \mathbf{y} distinct from $\mathbf{x}_1, \dots, \mathbf{x}_n$. If \mathbf{Z} is a formula, we say that $(S_{\mathbf{Y}_1 \dots \mathbf{Y}_n}^{\mathbf{x}_1 \dots \mathbf{x}_n}, \mathbf{Z})$ is the result of simultaneously substituting \mathbf{Y}_1 for $\mathbf{x}_1, \dots, \mathbf{x}_n$ in \mathbf{Z} .

Now that we have discussed the language of \mathcal{P} , let us make a few remarks on what we are about to do next, and illustrate them with Figure 1.7 In §11 we will explain the way this language is to be used by specifying certain axioms (initial principles) and rules of inference (rules of reasoning). We will thus have specified the *syntax* of the language. The syntactical aspects of a formal language are those aspects of it which can be discussed purely in

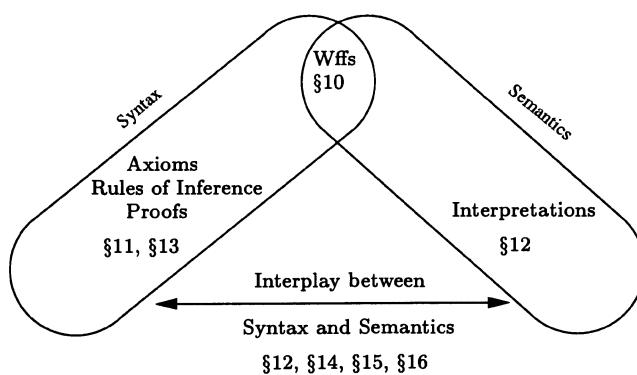


Figure 1.7: Plan for Chapter 1

terms of the symbols and sequences of symbols of the language, without any regard for what these might mean. In §12 we shall give a mathematically precise method of assigning meanings to the well-formed formulas, and thus specify the *semantics* of the language. After having specified the syntax and the semantics in ways which are quite independent of each other, except that they concern the same set of well-formed formulas, we will investigate how the syntax and semantics are related to each other. It will be seen that the study of the interplay between syntax and semantics is fundamental to the study of logical systems.

§10A. Supplement on Induction

In this supplement we shall discuss induction in a very informal way, and this discussion can be skipped by readers already familiar with induction and complete induction. In Chapter 6 we shall give a formal definition of the natural numbers, and prove the Principle of Mathematical Induction (Theorem 6102).

We may informally describe the set of natural numbers as the set which contains 0, 1, 2, 3, 4, 5, 6, etc. (The definition of \mathbb{N} in §60 shows that the meaning of the word “etc.” here can be made very explicit, but we shall not do that now.) One of the fundamental facts about the natural numbers is the Principle of Mathematical Induction (PMI), which says that for any property P of numbers, if 0 has property P , and $n + 1$ has property P

whenever n is a natural number which has property P , then every natural number has property P . If we abbreviate “ n has property P ” by “ Pn ”, “ \mathbf{A} and \mathbf{B} ” by “[$\mathbf{A} \wedge \mathbf{B}$]”, “if \mathbf{A} , then \mathbf{B} ” by “[$\mathbf{A} \supset \mathbf{B}$]”, “for all natural numbers n ” by “[$\forall n$ ”, and “for any property P ” by “[$\forall P$ ”, then we can use the following symbolic form of the Principle of Mathematical Induction:

$$\forall P[[P0 \wedge \forall n[Pn \supset P(n+1)]] \supset \forall nPn]$$

A related principle, called the Principle of Complete Induction (PCI), can be symbolized as follows:

$$\forall R[\forall n[(\forall j < n)Rj \supset Rn] \supset \forall nRn]$$

This can be paraphrased as saying that for any property R , if every natural number n has property R whenever all the natural numbers less than n have property R , then every natural number has property R .

By using some elementary facts about natural numbers we can derive each of these principles from the other. We shall indicate how this can be done, using symbolic notation for the sake of brevity.

First, assume that we have accepted PCI, and wish to justify PMI. Thus, we assume $P0 \wedge \forall n[Pn \supset P(n+1)]$, and try to prove $\forall nPn$. Let us try to show that $\forall n[(\forall j < n)Pj \supset Pn]$, so that we can use PCI. Let n be any natural number such that $(\forall j < n)Pj$. If $n = 0$, then we already know Pn , since we have assumed $P0$. Otherwise, $n = m + 1$ for some natural number m . We know Pm since $(\forall j < m + 1)Pj$, so $P(m + 1)$ follows from our assumption $\forall n[Pn \supset P(n+1)]$. Thus in each case we have established Pn , so since n was arbitrary, we have shown that $\forall n[(\forall j < n)Pj \supset Pn]$. Hence $\forall nPn$ by PCI, which is the desired conclusion.

Next we show how PCI can be derived from PMI. Assume R is a property such that $\forall n[(\forall j < n)Rj \supset Rn]$, and try to prove $\forall nRn$. In order to apply PMI, we shall let P be the property such that Pn means $(\forall j \leq n)Rj$, i.e., every natural number less than or equal to n has property R . We must prove $P0$ and $\forall n[Pn \supset P(n+1)]$. We are given that $(\forall j < 0)Rj \supset R0$, but there are no natural numbers less than 0, so $(\forall j < 0)Rj$ is vacuously true, so $R0$, and hence $P0$. Next assume Pn and try to derive $P(n+1)$. Pn means that $(\forall j \leq n)Rj$, and hence that $(\forall j < n+1)Rj$, and hence $R(n+1)$ by our assumption about R . From $(\forall j \leq n)Rj$ and $R(n+1)$ we infer $(\forall j \leq n+1)Rj$, which is $P(n+1)$. Thus from Pn we derive $P(n+1)$, so $\forall n[Pn \supset P(n+1)]$. Hence by PMI we obtain $\forall nPn$, which implies $\forall nRn$, and the proof of PCI is complete.

Finally, we show how metatheorem 1000 can be proved by complete induction and some basic facts about wffs. Assume that \mathcal{R} is a property of wffs with properties (1), (2), and (3) of metatheorem 1000. For any wff \mathbf{D} , let $\#\mathbf{D}$ (the length of \mathbf{D}) be the number of occurrences of primitive symbols in \mathbf{D} . Let P be the property of natural numbers such that P_n means $\forall \mathbf{D}[\#\mathbf{D} = n \supset \mathcal{R}(\mathbf{D})]$, i.e., every wff of length n has property \mathcal{R} . We must show that $\forall n[(\forall j < n)P_j \supset P_n]$, so let n be any natural number, and assume that $(\forall j < n)P_j$. To show P_n , let \mathbf{D} be any wff such that $\#\mathbf{D} = n$, and show $\mathcal{R}(\mathbf{D})$.

Now we use the facts that \mathbf{D} has one of the forms p , $\sim \mathbf{A}$, or $[\mathbf{A} \vee \mathbf{B}]$ (Exercise X1002), and that in the latter cases $\#\mathbf{A} < \#\mathbf{D}$ and $\#\mathbf{B} < \#\mathbf{D}$. If \mathbf{D} is p , then $\mathcal{R}(\mathbf{D})$ by (1). If \mathbf{D} is $\sim \mathbf{A}$, then from $(\forall j < n)P_j$ we infer P_j , where $j = \#\mathbf{A}$, so $\mathcal{R}(\mathbf{A})$ (by the definition of P), so $\mathcal{R}(\mathbf{D})$ by (2). Similarly, if \mathbf{D} is $[\mathbf{A} \vee \mathbf{B}]$, we infer $P(\#\mathbf{A})$ and $P(\#\mathbf{B})$, and obtain $\mathcal{R}(\mathbf{D})$ by (3). Thus in all cases we have shown $\mathcal{R}(\mathbf{D})$.

In summary, we have shown that $\forall \mathbf{D}[\#\mathbf{D} = n \supset \mathcal{R}(\mathbf{D})]$, which is P_n . Since n was arbitrary, we have shown that $\forall n[(\forall j < n)P_j \supset P_n]$, so $\forall n P_n$ follows by PCI. Now let \mathbf{D} be any wff, and let $n = \#\mathbf{D}$. We have established that P_n , so $\mathcal{R}(\mathbf{D})$ follows. This completes the proof of metatheorem 1000 using PCI.

EXERCISES

X1000. Using the definition of the set of wffs, show that:

- (a) Every formula consisting of a propositional variable standing alone is a wff.
- (b) If \mathbf{A} is a wff, then $\sim \mathbf{A}$ is a wff.
- (c) If \mathbf{A} and \mathbf{B} are wffs, then $[\mathbf{A} \vee \mathbf{B}]$ is a wff.

X1001. A *formation sequence for a formula \mathbf{Y}* is a finite sequence $\mathbf{Y}_1, \dots, \mathbf{Y}_m$ of formulas such that \mathbf{Y}_m is \mathbf{Y} and for each i ($1 \leq i \leq m$), one of the following conditions holds:

- (a) \mathbf{Y}_i is a propositional variable;
- (b) there is $j < i$ such that \mathbf{Y}_i is $\sim \mathbf{Y}_j$;
- (c) there are indices $j < i$ and $k < i$ such that \mathbf{Y}_i is $[\mathbf{Y}_j \vee \mathbf{Y}_k]$.

Prove that a formula \mathbf{Y} is well-formed iff it has a formation sequence.
(Hint: Metatheorem 1000 may be useful.)

X1002. Prove that every wff is of the form p for some propositional variable p , or of the form $\sim A$ or $[A \vee B]$ for some wffs A and B .

X1003. Using the results above, describe an effective procedure which can be used to determine of an arbitrary formula Y whether or not it is wf, and prove that this procedure is correct.

X1004. Decide whether the following formulas are wffs, and provide proofs to justify your answers:

- (a) $\sim p[\sim q]$
- (b) $[]p$
- (c) $[\sim p \vee]q$
- (d) $\sim [\sim \sim q \vee r]$

X1005. The following procedure is useful in checking whether complex expressions are wffs, or abbreviations of wffs. Given a formula A , consider the occurrences of symbols in A one-by-one in left-to-right order, and, starting with a count of zero, increase the count by one every time a [is encountered, and decrease the count by one every time a] is encountered.

EXAMPLE:

$$[[\sim r \vee \sim p] \vee \sim [[q \vee r] \vee [p \vee \sim s]] \vee q]]$$

1 2 1 2 3 4 3 4 3 2 10

Prove that if the expression is a wff, then the final count is zero, and the count never becomes negative. *Hint:* Use metatheorem 1000.

X1006. Show that if θ is any substitution for variables and A and B are wffs,

- (a) $(\theta \sim A) = \sim (\theta A)$
- (b) $(\theta[A \vee B]) = [(\theta A) \vee (\theta B)]$

X1007. Show that if p_1, \dots, p_n are distinct variables and B_1, \dots, B_n, A are wffs, then $(S_{B_1 \dots B_n}^{P_1 \dots P_n} A)$ is a wff. *Hint:* Use metatheorem 1000.

X1008. Evaluate $S_{[q \vee r]}^p \sim_p [\sim [\sim p \vee q] \vee \sim \sim p]$.

X1009. Show that if σ and τ are substitutions, then the composition $\sigma \circ \tau$ of σ with τ is also a substitution.

X1010. Express the statements below in symbolic notation. Let j mean “John can do the job”,

a mean “Alice can do the job”,
 w mean “Alice is willing to do the job”,
 g mean “Alice will get the job”,
 b mean “Bill will help John”,
 v mean “John will be willing to do the job”,
 d mean “John will do the job”.

- (a) If John can't do the job, then Alice can't either.
- (b) If Alice is willing and able to do the job, then she will get the job.
- (c) John can do the job only if Bill will help him.
- (d) If Bill will help John, then John will be willing to do the job.
- (e) John won't do the job unless Bill will help him.
- (f) John will do the job if and only if Bill will help him.

X1011. Prove that if $\mathbf{X}[\mathbf{Y}]\mathbf{Z}$ is a wff, where \mathbf{X} , \mathbf{Y} , and \mathbf{Z} are formulas of length ≥ 0 , then some propositional variable occurs in \mathbf{Y} . (*Hint:* Use metatheorem 1000 to prove some assertion which implies this.)

§11. The Axiomatic Structure of \mathcal{P}

Before discussing the details of the axiomatic structure of \mathcal{P} , let us make some general remarks about the axiomatic method and the role of syntax. When one uses the axiomatic method, one chooses as *axioms* certain statements whose consequences one wishes to investigate, and then derives their consequences (which one calls *theorems*) by the methods of logic. The set of theorems which can be derived from the axioms is called a *theory*, though the theory is sometimes identified with its set of axioms. It should be noted that people sometimes set up theories in order to explore ideas which are still in the process of being clarified. Thus, while the expressions used in the theory are not likely to be completely meaningless, the semantics of the theory may be vague or subject to revision. It is important to have a way of exploring the consequences of the axioms which depends only on syntax, and is not dependent on perhaps vague or shifting semantics.

While words may mean different things to different people, and it is sometimes difficult to know whether different people really attach the same meanings to various words and phrases, one way of understanding what people mean by words and phrases is to observe (over a long period) how they use these words and phrases in various situations. When one is concerned with the communication of mathematical ideas, which may be very abstract,

people may have different inner mental “images” of mathematical entities (including relations, functions, sets, etc.), but effective communication about these entities is still possible if there is agreement about how to use the language which describes these entities.

Thus, the syntactical aspect of a language plays an extremely important role even if people start out with an understanding of the language which is guided by some semantical perspective. The axiomatic method provides a way of specifying how a language describing some logically coherent discipline is to be *used* by giving rules for asserting certain statements.

Note the role of the axiomatic method in dealing with complexity. Many of the phenomena (including mathematical phenomena) which we would like to understand are extremely complicated, and even if one has a clear understanding of the meanings of the words in a sentence, one may not know whether or not the sentence is true. One needs to have a method of reasoning from simple statements whose truth is evident to complex statements whose truth may not be evident. The axiomatic method provides a way of doing this.

We shall illustrate how the axiomatic method works by introducing axioms and a rule of inference for the system \mathcal{P} , and deriving theorems which express general truths involving propositional connectives. The reader can probably think of a variety wffs which might be taken as axioms, such as :

- $p \vee \sim p$
- $p \supset p$
- $p \supset \neg p \vee q$
- $p \supset \neg q \vee p$
- $p \wedge q \supset p$
- $p \vee p \supset p$
- $q \vee p \supset p \vee q$
- $p \supset \neg \sim p \supset q$
- $[p \supset q] \wedge [q \supset r] \supset \neg p \supset r$

Actually, there are many different sets of axioms and rules of inference which one might reasonably choose, and each choice yields a different logistic system. Once one has chosen some axioms and specified one or more rules of

inference which can be used to derive theorems from the axioms, one may find that some of the axioms are derivable from others, and decide to modify the axiom set. Axiom systems tend to evolve as one learns more about them. The axioms and rule of inference for \mathcal{P} were chosen to be quite simple, independent (a matter which will be discussed in §13), and still adequate for quickly deriving the theorems that will be needed later.³

The axioms of \mathcal{P} consist of all wffs having one of the following three forms:

- (1) $\sim [A \vee A] \vee A$
- (2) $\sim A \vee \neg B \vee A$
- (3) $\sim [\sim A \vee B] \vee \neg \sim [C \vee A] \vee \neg B \vee C$

Note that \mathcal{P} has infinitely many axioms. We leave it to the reader to verify that there is an effective test for determining of any wff whether or not it is an axiom. We refer to the three forms of the axioms specified above as the axiom schemata for \mathcal{P} . Using the definition of \supset we may write them as follows:

Axiom Schema 1. $A \vee A \supset A$

Axiom Schema 2. $A \supset \neg B \vee A$

Axiom Schema 3. $A \supset B \supset \neg C \vee A \supset \neg B \vee C$

\mathcal{P} has just one rule of inference, which is:

Modus Ponens (MP). From A and $\sim A \vee B$, infer B .

Using the definition of \supset , we can also state Modus Ponens as follows: From A and $A \supset B$, infer B .

The rationale for using Modus Ponens as a rule of inference is provided by the meanings of the propositional connectives. In every situation where the statements A and $A \supset B$ are true, B must also be true.

We shall use H (with or without subscripts) as a syntactical variable for a set of wffs. An especially important case is that in which H denotes the empty set of wffs.

³ \mathcal{P} is very closely related to the system discussed in [Rasiowa, 1949].

DEFINITIONS.

A proof of a wff \mathbf{B} from the set \mathcal{H} of hypotheses is a finite sequence $\mathbf{B}_1, \dots, \mathbf{B}_m$ of wffs such that \mathbf{B}_m is \mathbf{B} and for each j ($1 \leq j \leq m$) at least one of the following conditions is satisfied:

- (1) \mathbf{B}_j is an axiom.
- (2) \mathbf{B}_j is a member of \mathcal{H} .
- (3) \mathbf{B}_j is inferred by Modus Ponens from wffs \mathbf{B}_i and \mathbf{B}_k , where $i < j$ and $k < j$.

An alternative way of expressing condition (3) is to say that there exist $i < j$ and $k < j$ such that \mathbf{B}_k is $[\mathbf{B}_i \supset \mathbf{B}_j]$.

A proof of a wff \mathbf{B} is a proof of \mathbf{B} from the empty set of hypotheses. A *theorem* is a wff which has a proof.

We write $\mathcal{H} \vdash \mathbf{B}$ to indicate that there is a proof of \mathbf{B} from \mathcal{H} and say that \mathbf{B} is *derivable from \mathcal{H}* . In place of $\{\mathbf{A}_1, \dots, \mathbf{A}_n\} \vdash \mathbf{B}$ we write $\mathbf{A}_1, \dots, \mathbf{A}_n \vdash \mathbf{B}$, and in place of $\mathcal{H} \cup \{\mathbf{A}\} \vdash \mathbf{B}$ we write $\mathcal{H}, \mathbf{A} \vdash \mathbf{B}$. We write $\vdash \mathbf{B}$ to indicate that \mathbf{B} is derivable from the empty set, i.e., that \mathbf{B} is a theorem. In order to avoid confusion when several logistic systems are under consideration, we may write $\vdash_{\mathcal{P}} \mathbf{B}$ (instead of $\vdash \mathbf{B}$) to indicate that \mathbf{B} is a theorem of \mathcal{P} .

We have now completed the specification of \mathcal{P} as a *logistic system*. To define a logistic system, one must specify its primitive symbols, formation rules, axioms, and rules of inference. A logistic system must be defined in such a way that there is an effective test (algorithm) for determining of any formula whether or not it is in fact a wff.⁴ A *proof* in a logistic system \mathcal{L} of a wff \mathbf{B} is a sequence of wffs, each of which is an axiom of \mathcal{L} or is inferred from preceding wffs in the sequence by a rule of inference of \mathcal{L} , with \mathbf{B} as the last wff in the sequence. A *theorem* of \mathcal{L} is a wff which has a proof in \mathcal{L} . It is part of the definition of a logistic system that there must be an effective test for determining of a sequence of formulas whether or not it is a proof.⁵ (The reader is asked to show this for the system \mathcal{P} in Exercise X1100.) The rationale for this is clear: a proof should carry conviction to one who accepts the axioms and the methods of reasoning embodied in the rules of inference. However, the basic purpose of a proof is defeated if one cannot always determine whether an alleged proof is actually a proof. As a matter of fact, mathematicians sometimes do offer proofs which their colleagues do not find

⁴The notion of effectiveness will be discussed in detail in §65.

⁵For more discussion of these matters the reader is referred to the Introduction of [Church 1956].

convincing. Occasionally such proofs are actually incorrect, but more often they are just so abbreviated that they are hard to follow. The idea of a logistic system is to provide a standard of rigor which will guarantee that when a proof is presented within a logistic system, one can verify that it is in fact a proof, and (if so) be confident that the theorem is really proved. While mathematical proofs are rarely presented with all the details that strict conformity to the definition of a proof in a logistic system demands, once one has an appropriate logistic system in mind, one can regard a mathematical proof as an abbreviated proof in the logistic system, and provide more details (thus making the proof less abbreviated) whenever questions about it arise. In principle, sufficient repetition of this process would eventually produce an unabbreviated proof in the logistic system. We can anticipate the development of automated theorem proving systems which will provide these additional details automatically, thus verifying the correctness of proofs and permitting readers to examine the proofs at whatever level of detail they find most congenial. (For convenience, we have referred to “mathematical proofs” in the discussion above, but the same considerations apply to proofs in other disciplines which use logic.)

While \mathcal{P} is only one example of a logistic system, the methods and concepts we shall use while investigating \mathcal{P} will be relevant to a variety of logistic systems.

One of the surprising things about many logistic systems is how much can be derived from simple axioms. This is also one of the reasons such systems are so useful; they provide an extremely economical way of storing a great wealth of information. We next explore the power of \mathcal{P} by proving certain theorems of \mathcal{P} and metatheorems about \mathcal{P} .

1100. If $\mathcal{H}_1 \vdash \mathbf{A}$ and $\mathcal{H}_1 \subseteq \mathcal{H}_2$, then $\mathcal{H}_2 \vdash \mathbf{A}$.

Proof: Any proof of \mathbf{A} from \mathcal{H}_1 must also be a proof of \mathbf{A} from \mathcal{H}_2 . ■

Corollary. If $\mathcal{H}_1 \vdash \mathbf{A}$ and $\mathcal{H}_2 \vdash \mathbf{A} \supset \mathbf{B}$ and $\mathcal{H}_1 \subseteq \mathcal{H}$ and $\mathcal{H}_2 \subseteq \mathcal{H}$, then $\mathcal{H} \vdash \mathbf{B}$.

REMARK. Metatheorem 1100 tells us that when using the system \mathcal{P} we can enlarge the set of hypotheses (which we can think of as expressing the information we consider) without losing any consequences we have derived from these hypotheses. Logistic systems which have this property are called *monotonic*. If one wishes to formalize reasoning in a context where additional information may lead one to change one’s conclusions, one needs to

use a *nonmonotonic* logistic system. We shall not study nonmonotonic logics in this book.

1101 Rule of Substitution (Sub). If $\mathcal{H} \vdash A$, and if p_1, \dots, p_n are distinct variables which do not occur in any wff in \mathcal{H} , then $\mathcal{H} \vdash S_{B_1 \dots B_n}^{p_1 \dots p_n} A$.

Proof: Let θ be the substitution $S_{B_1 \dots B_n}^{p_1 \dots p_n}$. It is easy to see that if C_1, \dots, C_m is a proof of A from \mathcal{H} , then $\theta C_1, \dots, \theta C_m$ is a proof of θA from \mathcal{H} . Note how the condition on the variables comes into play when C_i is a member of \mathcal{H} . ■

Of course, the sentence above is really a rather brief sketch of a proof of this metatheorem. For a more complete proof, show by complete induction on i that $\mathcal{H} \vdash \theta C_i$ for each i with $1 \leq i \leq m$. Break the proof into cases according to how C_i was justified in the original proof.

In proving theorems of \mathcal{P} , we shall indicate at the right-hand margin how each step of the proof is obtained. However, we shall often omit explicit mention of the rule of substitution and metatheorem 1100. We shall also sometimes combine several steps into one.

1102. If $\mathcal{H} \vdash A \supset B$ and $\mathcal{H} \vdash C \vee A$, then $\mathcal{H} \vdash B \vee C$.

Proof: Suppose we are given proofs D_1, \dots, D_i of $A \supset B$ and E_1, \dots, E_j of $C \vee A$ from the set \mathcal{H} of hypotheses. Thus D_i is $[A \supset B]$ and E_j is $[C \vee A]$. It is easy to see that the following is a proof from \mathcal{H} of $B \vee C$:

(1)	D_1	
	⋮	
(i)	D_i	
(i + 1)	$A \supset B \supset C \vee A \supset B \vee C$	Axiom 3
(i + 2)	$C \vee A \supset B \vee C$	MP: $i, i + 1$
(i + 3)	E_1	
	⋮	
(i + j + 2)	E_j	
(i + j + 3)	$B \vee C$	MP: $i + 2, i + j + 2$

■

Henceforth, we shall abbreviate proofs such as that above, and simply say something like “by Axiom 3 and MP” to indicate the main ideas used in the proof.

Note that metatheorem 1102 can be used as a rule of inference, since it shows that one can infer $\mathbf{B} \vee \mathbf{C}$ from $\mathbf{A} \supset \mathbf{B}$ and $\mathbf{C} \vee \mathbf{A}$. Moreover, the proof of 1102 shows explicitly how to construct a proof of $\mathbf{B} \vee \mathbf{C}$ if one is given proofs of $\mathbf{A} \supset \mathbf{B}$ and $\mathbf{C} \vee \mathbf{A}$. Metatheorems such as 1102 which can be used as rules of inference are called *derived rules of inference*. For contrast, we refer to Modus Ponens as the *primitive* rule of inference of \mathcal{P} .

1103. $\vdash p \vee \sim p$

Proof:

$$(.) \vdash p \vee p \supset p \supset \neg \sim p \vee [p \vee p] \supset \neg p \vee \sim p \quad \text{Axiom 3}$$

$$(.) \vdash p \vee \sim p \quad \text{MP: Axiom 1, .1, Axiom 2}$$

■

1104. $\vdash p \supset \sim \sim p$ Sub: 1103

■

1105. $\vdash \sim \sim p \supset p$

Proof:

$$(.) \vdash \neg p \supset \sim \sim p \supset \neg p \vee \sim p \supset \neg \sim \sim p \vee p \quad \text{Axiom 3}$$

$$(.) \vdash \neg \sim p \supset p \quad \text{Sub: 1104; MP: .1; MP: 1103}$$

■

1106. $\vdash p \supset p$

REMARK. The proof below is quite easy to understand if one has a good intuitive understanding of 1102. When one has proved $\mathbf{A} \supset \mathbf{B}$, one can use 1102 to

replace	$\mathbf{C} \vee \mathbf{A}$
by	$\mathbf{B} \vee \mathbf{C}$.

In the proof below, (.1) and (.2) are preliminary results. The essential part of the proof consists of systematically transforming Axiom 2 in line (.3) into line (.8), from which (.11) is easily obtained.

Proof of 1106:

(.1)	\vdash	$p \vee p \supset \neg\neg p \vee p$	MP: 1104, Axiom 3
(.2)	\vdash	$\neg\neg p \vee p \supset \neg\neg p \vee \neg\neg p$	MP: 1104, Axiom 3
(.3)	\vdash	$\neg p \vee [p \vee p]$	Axiom 2
(.4)	\vdash	$[\neg\neg p \vee p] \vee \neg p$	1102: .1, .3
(.5)	\vdash	$\neg\neg\neg p \vee [\neg\neg p \vee p]$	Sub: 1104; 1102: .4
(.6)	\vdash	$[\neg\neg p \vee \neg\neg p] \vee \neg\neg\neg p$	1102: .2, .5
(.7)	\vdash	$\neg p \vee [\neg\neg p \vee \neg\neg p]$	Sub: 1105; 1102: .6
(.8)	\vdash	$\neg\neg p \vee \neg p$	1102: Axiom 1, .7
(.9)	\vdash	$\neg\neg\neg p \vee \neg p$	Sub: .8
(.10)	\vdash	$p \vee \neg\neg\neg p$	1102 : 1105, .9
(.11)	\vdash	$\neg p \vee p$	Sub: 1105; 1102: .10

■

In order to see the usefulness of derived rules of inference, the reader is invited to write out a complete proof of 1106 using only the primitive rule of inference, and compare it with the abbreviated proof using derived rules of inference which is given above.

$$\mathbf{1107.} \vdash q \vee p \supset p \vee q \quad \text{MP: 1106, Axiom 3}$$

■

$$\mathbf{1108.} \text{ If } \mathcal{H} \vdash \mathbf{A} \vee \mathbf{B}, \text{ then } \mathcal{H} \vdash \mathbf{B} \vee \mathbf{A}.$$

Proof: By 1107. ■

1109 Transitive Law of Implication (Trans). If $\mathcal{H} \vdash \mathbf{A}_1 \supset \mathbf{A}_2$, $\mathcal{H} \vdash \mathbf{A}_2 \supset \mathbf{A}_3, \dots$, and $\mathcal{H} \vdash \mathbf{A}_{n-1} \supset \mathbf{A}_n$, then $\mathcal{H} \vdash \mathbf{A}_1 \supset \mathbf{A}_n$.

Proof: It suffices to consider the case in which $n = 3$, since the general case follows easily from this one by mathematical induction.

(.1)	$\mathcal{H} \vdash \neg\mathbf{A}_1 \vee \mathbf{A}_2$	given
(.2)	$\mathcal{H} \vdash \mathbf{A}_2 \supset \mathbf{A}_3$	given
(.3)	$\mathcal{H} \vdash \mathbf{A}_3 \vee \neg\mathbf{A}_1$	1102: .2, .1
(.4)	$\mathcal{H} \vdash \mathbf{A}_1 \supset \mathbf{A}_3$	1108: .3

■

1110. If $\mathcal{H} \vdash A \supset C$ and $\mathcal{H} \vdash B \supset C$, then $\mathcal{H} \vdash A \vee B \supset C$.

Proof:

(.1) $\mathcal{H} \vdash B \supset C$	given
(.2) $\mathcal{H} \vdash A \vee B \supset C \vee A$	MP: .1, Axiom 3
(.3) $\mathcal{H} \vdash A \supset C$	given
(.4) $\mathcal{H} \vdash C \vee A \supset C \vee C$	MP: .3, Axiom 3
(.5) $\mathcal{H} \vdash C \vee C \supset C$	Axiom 1
(.6) $\mathcal{H} \vdash A \vee B \supset C$	Trans: .2, .4, .5

■

1111. If $\mathcal{H} \vdash A \supset C$ and $\mathcal{H} \vdash \sim A \supset C$, then $\mathcal{H} \vdash C$.

Proof: By 1110 and 1103. ■

1112. If $\mathcal{H} \vdash A \supset B$, then $\mathcal{H} \vdash A \supset \neg C \vee B$ and $\mathcal{H} \vdash A \supset \neg B \vee C$.

Proof:

(.1) $\mathcal{H} \vdash A \supset B$	given
(.2) $\mathcal{H} \vdash A \supset \neg C \vee B$	Trans: .1, Axiom 2
(.3) $\mathcal{H} \vdash A \supset \neg B \vee C$	Sub: 1107; Trans: .2

■

1113. $\vdash [p \vee q] \vee r \supset \neg p \vee \neg q \vee r$

Proof:

(.1) $\vdash p \supset \neg p \vee \neg q \vee r$	1112: 1106
(.2) $\vdash q \supset \neg q \vee r$	1112: 1106
(.3) $\vdash q \supset \neg p \vee \neg q \vee r$	1112: .2
(.4) $\vdash p \vee q \supset \neg p \vee \neg q \vee r$	1110: .1, .3
(.5) $\vdash r \supset \neg p \vee \neg q \vee r$	1112: 1106
(.6) $\vdash [p \vee q] \vee r \supset \neg p \vee \neg q \vee r$	1110: .4, .5

■

1114. If $\mathcal{H} \vdash [A \vee B] \vee C$, then $\mathcal{H} \vdash A \vee \neg B \vee C$.

Proof: By 1113. ■

REMARK. The converse of 1114 is established in exercise X1101.

1115. If $\mathcal{H} \vdash A \supset B$ and $\mathcal{H} \vdash A \supset \neg B \supset C$, then $\mathcal{H} \vdash A \supset C$.

Proof:

- | | |
|--|-------------------|
| (.1) $\mathcal{H} \vdash A \supset B$ | given |
| (.2) $\mathcal{H} \vdash \neg A \vee \neg \neg B \vee C$ | given |
| (.3) $\mathcal{H} \vdash [\neg B \vee C] \vee \neg A$ | 1108: .2 |
| (.4) $\mathcal{H} \vdash B \supset \neg C \vee \neg A$ | 1114: .3 |
| (.5) $\mathcal{H} \vdash A \supset \neg C \vee \neg A$ | Trans: .1, .4 |
| (.6) $\mathcal{H} \vdash C \vee \neg A \vee \neg A$ | 1108: .5 |
| (.7) $\mathcal{H} \vdash C \vee \neg \neg A \vee \neg A$ | 1114: .6 |
| (.8) $\mathcal{H} \vdash \neg A \vee C$ | 1102: Axiom 1, .7 |

■

1116. Deduction Theorem. If $\mathcal{H}, A \vdash B$, then $\mathcal{H} \vdash A \supset B$.

Proof: Let C_1, \dots, C_m be a proof of B from the hypotheses \mathcal{H}, A . We prove $\mathcal{H} \vdash A \supset C_i$ for all i ($1 \leq i \leq m$) by complete induction on i .

Case 1. C_i is an axiom or a member of \mathcal{H} .

- | | |
|---|---------------------------------|
| (.11) $\mathcal{H} \vdash C_i$ | C_i is an axiom or hypothesis |
| (.12) $\mathcal{H} \vdash C_i \supset \neg A \supset C_i$ | Axiom 2 |
| (.13) $\mathcal{H} \vdash A \supset C_i$ | MP: .11, .12 |

Case 2. C_i is A .

- | | |
|--|-----------|
| (.21) $\vdash A \supset C_i$ | Sub: 1106 |
| (.22) $\mathcal{H} \vdash A \supset C_i$ | 1100: .21 |

Case 3. In the proof from $\mathcal{H} \cup \{ A \}$, C_i is inferred by Modus Ponens from C_j and C_k , where $j < i$, $k < i$, and C_k is $[C_j \supset C_i]$.

- | | |
|--|-------------------------|
| (.31) $\mathcal{H} \vdash A \supset C_j$ | by inductive hypothesis |
| (.32) $\mathcal{H} \vdash A \supset [C_j \supset C_i]$ | by inductive hypothesis |
| (.33) $\mathcal{H} \vdash A \supset C_i$ | 1115: .31, .32 |

Since by the definition of a proof from hypotheses C_i must fall under one of these cases, the inductive proof is complete. Thus $\mathcal{H} \vdash A \supset C_m$, and C_m is B , so $\mathcal{H} \vdash A \supset B$. ■

1117. $\vdash \sim p \supset \neg \sim q \supset \sim \neg p \vee q$

Proof:

(.1) $\vdash p \supset \neg \sim p \vee \sim \sim q$	1112: 1104
(.2) $\vdash q \supset \neg \sim p \vee \sim \sim q$	1112: 1104
(.3) $\vdash p \vee q \supset \neg \sim p \vee \sim \sim q$	1110: .1, .2
(.4) $\sim \sim p \vee \sim \sim q \vee \neg \sim p \vee q$	1108: .3
(.5) $\sim \sim p \vee \neg \sim q \vee \neg \sim p \vee q$	1114: .4

■

1118. $\vdash p \supset \neg \sim p \supset q$ 1112: 1104 ■

DEFINITION. Let \mathcal{R} be any property of wffs. We say that a rule of inference *preserves* property \mathcal{R} iff a wff inferred by the rule from wffs which have the property must also have the property.

1119. Principle of Induction on Proofs. Let \mathcal{L} be any logistic system, and let \mathcal{R} be any property of wffs of \mathcal{L} . If every axiom of \mathcal{L} has property \mathcal{R} , and each rule of inference of \mathcal{L} preserves property \mathcal{R} , then every theorem of \mathcal{L} has property \mathcal{R} .

Proof: Let C_1, \dots, C_m be a proof of B in \mathcal{L} . It is easy to establish that C_i has property \mathcal{R} for all i ($1 \leq i \leq m$) by complete induction on i . Either C_i is an axiom, and therefore has property \mathcal{R} , or C_i is inferred by a rule of inference from preceding wffs in the proof. In the latter case the preceding wffs have property \mathcal{R} by inductive hypothesis, so C_i has the property since the rules of inference preserve the property. In particular, C_m , which is B , must have property \mathcal{R} . ■

EXERCISES

X1100. Show that there is an effective test for determining whether an arbitrary finite sequence of formulas is a proof.

X1101. Show that in any system of logic in which 1108 and 1114 are primitive or derived rules of inference, from $A \vee \neg B \vee C$ one may infer $[A \vee B] \vee C$.

X1102. Let \mathcal{L} be a formulation of propositional calculus in which the sole connectives are negation and disjunction, the sole rule of inference is Modus

Ponens, and with the following axiom schema:

$$\sim [A \vee B] \vee [B \vee C]$$

Show that every wff of \mathcal{L} is a theorem of \mathcal{L} . (Such a logistic system is said to be *absolutely inconsistent*.) (*Hint*: You should think of this as a purely syntactic problem involving symbol manipulation, since \mathcal{L} obviously does not make sense when the connectives are interpreted in the usual way. It may be helpful to prove some derived rules of inference for \mathcal{L} .)

X1103. Let \mathcal{M} be the system which has the same wffs and rule of inference as \mathcal{P} , and the single axiom schema

$$[A \supset B \supset \neg C \vee \neg D \vee E] \supset \neg D \supset A \supset \neg C \vee \neg E \vee A.$$

Show that each theorem of \mathcal{P} is a theorem of \mathcal{M} . (For help see [Meredith, 1953].)

§12. Semantics, Consistency, and Completeness of \mathcal{P}

In §§10 – 11 we have been concerned with the syntax of \mathcal{P} . In this section we shall define the semantics of \mathcal{P} in a mathematically precise way, and discuss some important aspects of the relation between the syntax and the semantics of \mathcal{P} .

We said earlier that the propositional variables of \mathcal{P} would play the role of statements. However, in any particular context, a statement is either true or false, so in defining the formal semantics of \mathcal{P} we shall simply assign truth values to these variables.

DEFINITION. An *assignment* of truth values to propositional variables is a function from the set of variables to the set $\{\text{T}, \text{F}\}$ of truth values. (We use T as an abbreviation for Truth, and F for Falsehood.) Occasionally we shall also consider assignments which have as domain some subset of the set of all variables.

Once one has specified the values of the variables in a wff, the wff itself becomes either true or false, and we say that its *value* is truth or falsehood. This is made precise in the following definition.

DEFINITION. The *value* $\mathcal{V}_\varphi A$ of a wff A with respect to the assignment φ is defined as follows by induction on the construction of A :

- (1) $\mathcal{V}_\varphi \mathbf{p} = \varphi \mathbf{p}$
- (2) $\mathcal{V}_\varphi \sim \mathbf{B} = \begin{cases} \mathbf{T} & \text{if } \mathcal{V}_\varphi \mathbf{B} = \mathbf{F} \\ \mathbf{F} & \text{if } \mathcal{V}_\varphi \mathbf{B} = \mathbf{T} \end{cases}$
- (3) $\mathcal{V}_\varphi [\mathbf{B} \vee \mathbf{C}] = \begin{cases} \mathbf{T} & \text{if } \mathcal{V}_\varphi \mathbf{B} = \mathbf{T} \text{ or } \mathcal{V}_\varphi \mathbf{C} = \mathbf{T} \\ \mathbf{F} & \text{if } \mathcal{V}_\varphi \mathbf{B} = \mathbf{F} \text{ and } \mathcal{V}_\varphi \mathbf{C} = \mathbf{F}. \end{cases}$

We say that \mathbf{A} is *true* with respect to φ if $\mathcal{V}_\varphi \mathbf{A} = \mathbf{T}$, and *false* with respect to φ if $\mathcal{V}_\varphi \mathbf{A} = \mathbf{F}$.

Since we shall have many occasions to use the definition of \mathcal{V}_φ in proofs, it is convenient to extend our meta-language so that the definition can be expressed more concisely. To this end, let us introduce \sim and \vee into our meta-language as operators on truth values defined as follows: $\sim \mathbf{T} = \mathbf{F}$; $\sim \mathbf{F} = \mathbf{T}$; $\mathbf{T} \vee \mathbf{T} = \mathbf{T}$; $\mathbf{T} \vee \mathbf{F} = \mathbf{T}$; $\mathbf{F} \vee \mathbf{T} = \mathbf{T}$; $\mathbf{F} \vee \mathbf{F} = \mathbf{F}$. Thus, as symbols of the meta-language \vee and \sim now have the same meaning as we have just given them in the object language.

Of course, we thus introduce some ambiguity into the meta-language, since \sim and \vee also serve in the meta-language as names for themselves as symbols of the object language. This ambiguity could be obviated by using different symbols in the meta-language, or by coloring symbols a different color when they are used as symbols of the meta-language, and we leave it to the reader to mentally color the symbols of this text in an appropriate way.

Using the notation just introduced, we can express the inductive definition of \mathcal{V}_φ as follows:

- (1) $\mathcal{V}_\varphi \mathbf{p} = \varphi \mathbf{p}$
- (2) $\mathcal{V}_\varphi \sim \mathbf{B} = \sim \mathcal{V}_\varphi \mathbf{B}$
- (3) $\mathcal{V}_\varphi [\mathbf{B} \vee \mathbf{C}] = (\mathcal{V}_\varphi \mathbf{B}) \vee (\mathcal{V}_\varphi \mathbf{C}).$

The reader should take care to note how each occurrence of \sim or \vee is being used here, and provide appropriate mental colorings of these statements.

DEFINITIONS.

- (1) \mathbf{A} is a *tautology* iff $\mathcal{V}_\varphi \mathbf{A} = \mathbf{T}$ for all assignments φ . We write $\models \mathbf{A}$ to indicate that \mathbf{A} is a tautology.
- (2) \mathbf{A} is a *contradiction* iff $\mathcal{V}_\varphi \mathbf{A} = \mathbf{F}$ for all assignments φ .
- (3) \mathbf{A} implies \mathbf{B} (by propositional calculus) iff $[\mathbf{A} \supset \mathbf{B}]$ is a tautology.
- (4) \mathbf{A} is equivalent to \mathbf{B} (by propositional calculus) iff $[\mathbf{A} \equiv \mathbf{B}]$ is a tautology. Thus, \mathbf{A} is equivalent to \mathbf{B} iff $\mathcal{V}_\varphi \mathbf{A} = \mathcal{V}_\varphi \mathbf{B}$ for all assignments φ .

p	q	$p \wedge q$	$\supset p$	$p \vee q$
T	T	T	T	T
T	F	F	T	T
F	T	F	T	T
F	F	F	T	F

Figure 1.8: A simple tautology

- (5) An assignment φ *falsifies* \mathbf{A} iff $\mathcal{V}_\varphi \mathbf{A} = \mathbf{F}$.
- (6) An assignment φ *satisfies* \mathbf{A} iff $\mathcal{V}_\varphi \mathbf{A} = \mathbf{T}$. We write $\models_\varphi \mathbf{A}$ to indicate that φ satisfies \mathbf{A} . An alternative notation for this is $\models \mathbf{A}[\varphi]$.
- (7) An assignment φ *satisfies* a set \mathcal{S} of wffs iff φ satisfies every wff of \mathcal{S} . (Every assignment satisfies the empty set of wffs.)
- (8) A set \mathcal{S} of wffs is *satisfiable* iff there is an assignment which satisfies it. The set is *unsatisfiable* or *contradictory* iff it is not satisfiable.

In a truth table, each horizontal line corresponds to some assignment. Thus a wff is a tautology iff the column under its main connective contains only T's (as in Figure 1.8, for example).

It is easy to see that if n propositional variables occur in a wff A , then there are 2^n possible assignments of values to the variables of \mathbf{A} , since each variable can be assigned either of two values. Wffs with many variables have very large truth tables, so it is nice to be able to show that a wff is a tautology without writing out the entire truth table. A method which enables one to do this for some wffs is illustrated in Figure 1.9. One wishes to investigate whether $p \wedge [p \supset q] \supset q \vee r$ is a tautology, so one supposes (1) that there is an assignment with respect to which it is false. One then considers what the values of various wf parts of the given wff must be with respect to this assignment. Since $[A \supset B]$ can be false only when \mathbf{A} is true and \mathbf{B} is false, (2) $p \wedge [p \supset q]$ must be true, while (3) $q \vee r$ must be false. Since $[A \wedge B]$ is true only if \mathbf{A} and \mathbf{B} are both true, (4) p must be true, and (5) $p \supset q$ must be true, so (6) q must be true. This makes $q \vee r$ true (regardless of the value of r), which contradicts (3). This contradiction shows that there is no assignment which falsifies the wff, so it must be a tautology.

Of course, when one applies the method above to certain examples, one may eventually reach a point where one must consider several possible cases, but one still avoids writing out the entire truth table. If one applies the

method to a wff which is actually not a tautology, one will obtain an assignment which falsifies the wff, and thus see that it is not a tautology.

DEFINITION. We say that a wff of a logistic system is *valid* with respect to an interpretation iff the value of that wff is truth (under that interpretation) for all assignments of values to its (free)⁶ variables. We shall write $\models \mathbf{A}$ to indicate that \mathbf{A} is valid. An interpretation of a logistic system is *sound* iff, under that interpretation, the axioms are all valid and the rules of inference preserve validity. Clearly, in this case the theorems must all be valid.

Note that a wff of \mathcal{P} is a tautology iff it is valid under the interpretation which is implicit in the definition of the valuation function \mathcal{V}_φ .

1200 Soundness Theorem. Every theorem of \mathcal{P} is a tautology.

Proof: It is easy to check that each axiom is a tautology. Also, if $\mathcal{V}_\varphi \mathbf{A} = \top$ and $\mathcal{V}_\varphi [\sim \mathbf{A} \vee \mathbf{B}] = \top$ then $\mathcal{V}_\varphi \mathbf{B} = \top$. Hence if \mathbf{A} and $[\sim \mathbf{A} \vee \mathbf{B}]$ are tautologies, \mathbf{B} must also be a tautology. Therefore we see by 1119 that every theorem is a tautology. ■

If every wff is a theorem of a particular logistic system, then a wff gains no significance by being a theorem of that system. Such systems are called *inconsistent*. Actually, this notion can be defined in several ways.

DEFINITION. A logistic system is *absolutely inconsistent* iff every wff is a theorem. A logistic system is *inconsistent with respect to negation* iff there is a wff \mathbf{A} such that both \mathbf{A} and the negation of \mathbf{A} are theorems. A logistic system is *absolutely consistent* iff it is not absolutely inconsistent, and it is *consistent with respect to negation* iff it is not inconsistent with respect to negation. A set \mathcal{H} of wffs is *consistent* iff there is a wff \mathbf{A} such that not $\mathcal{H} \vdash \mathbf{A}$.

⁶In propositional calculus, all occurrences of variables are free. We insert the qualifying adjective “free” here so that this definition will be applicable to systems which we shall consider later.

p	\wedge	$[p \supset q]$	\supset	q	\vee	r
T	T	T	T	T	F	
4	2	4	5	6	1	6

x

Figure 1.9: A short method of verifying a tautology

By using 1118 and 1101 it is easy to see that every wff of \mathcal{P} of the form $[A \supset \neg A \supset B]$ is a theorem of \mathcal{P} . The expression $[A \supset \neg A \supset B]$, which contains syntactical variables and belongs to the meta-language, is called a *theorem schema* of \mathcal{P} .

1201 Theorem. Let \mathcal{S} be any logistic system in which $A \supset \neg A \supset B$ is a theorem schema and in which Modus Ponens is a primitive or derived rule of inference. Then \mathcal{S} is absolutely consistent iff \mathcal{S} is consistent with respect to negation.

Proof: If \mathcal{S} is inconsistent with respect to negation there is a wff A such that $\vdash A$ and $\vdash \neg A$. Hence by two applications of Modus Ponens to $[A \supset \neg A \supset B]$ we obtain $\vdash B$ for any wff B , so \mathcal{S} is absolutely inconsistent. Conversely, if \mathcal{S} is absolutely inconsistent, then $\vdash A$ and $\vdash \neg A$ for every wff, so \mathcal{S} is inconsistent with respect to negation.

1202 Consistency Theorem. \mathcal{P} is absolutely consistent and consistent with respect to negation.

Proof: If A is any theorem of \mathcal{P} , then A is a tautology, so $\neg A$ is not a tautology, so $\neg A$ is not a theorem. Thus \mathcal{P} is consistent with respect to negation, and hence, by 1118, 1101, and 1201, absolutely consistent. ■

Next we prepare to prove the converse of 1200.

1203 Lemma. Let A be a wff and let all the variables in A be among p_1, \dots, p_n . Let φ be any assignment. For any wff B , let

$$\begin{aligned} B^\varphi &= B \text{ if } V_\varphi B = T \\ &= \neg B \text{ if } V_\varphi B = F. \end{aligned}$$

Then $p_1^\varphi, \dots, p_n^\varphi \vdash A^\varphi$.

EXAMPLE: $p, \neg q, r \vdash \neg p \supset q \wedge r$

Proof: The proof is by induction on the construction of A .

Case 1. A is p_i . Then $p_1^\varphi, \dots, p_n^\varphi \vdash p_i^\varphi$.

Case 2a. A is $\neg B$ and $V_\varphi B = T$.

$$\begin{aligned} p_1^\varphi, \dots, p_n^\varphi \vdash B \\ p_1^\varphi, \dots, p_n^\varphi \vdash B \supset \neg \neg B \end{aligned}$$

by inductive hypothesis
Sub: 1104

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \sim\sim \mathbf{B}$$

MP

Since \mathbf{A}^φ is $\sim\sim \mathbf{B}$, this is the desired conclusion.

Case 2b. \mathbf{A} is $\sim \mathbf{B}$ and $\mathcal{V}_\varphi \mathbf{B} = \mathsf{F}$.

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \sim \mathbf{B}$$

by inductive hypothesis

But \mathbf{A}^φ is $\sim \mathbf{B}$, so this is the desired conclusion.

Case 3a. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$ and $\mathcal{V}_\varphi \mathbf{B} = \mathsf{T}$.

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \mathbf{B}$$

by inductive hypothesis

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \mathbf{B} \supset \bullet \mathbf{B} \vee \mathbf{C}$$

1112: 1106

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \mathbf{B} \vee \mathbf{C}$$

MP

Case 3b. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$ and $\mathcal{V}_\varphi \mathbf{C} = \mathsf{T}$.

The proof proceeds as in Case 3a.

Case 3c. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$ and $\mathcal{V}_\varphi \mathbf{B} = \mathsf{F}$ and $\mathcal{V}_\varphi \mathbf{C} = \mathsf{F}$.

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \sim \mathbf{B}$$

by inductive hypothesis

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \sim \mathbf{C}$$

by inductive hypothesis

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \sim \mathbf{B} \supset \bullet \sim \mathbf{C} \supset \sim \bullet \mathbf{B} \vee \mathbf{C}$$

Sub: 1117

$$\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \sim \bullet \mathbf{B} \vee \mathbf{C}$$

MP

■

1204 Completeness Theorem. Every tautology is a theorem of \mathcal{P} .

Proof: Let \mathbf{A} be a tautology and let n be the number of variables which occur in \mathbf{A} . Let $\{\mathbf{p}_1, \dots, \mathbf{p}_j\}$ (where $0 \leq j \leq n$) be a set of distinct variables which occur in \mathbf{A} . We prove that $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \mathbf{A}$ for any assignment φ by induction on $n - j$.

For $n - j = 0$ this follows from 1203.

For the induction step, we consider assignments φ' and φ'' such that $\varphi' \mathbf{p}_i = \varphi \mathbf{p}_i = \varphi'' \mathbf{p}_i$ for $i \leq j$ and $\varphi' \mathbf{p}_{j+1} = \mathsf{T}$ and $\varphi'' \mathbf{p}_{j+1} = \mathsf{F}$ to see that:

- (.1) $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_j^\varphi, \mathbf{p}_{j+1} \vdash \mathbf{A}$ by inductive hypothesis (using φ')
- (.2) $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_j^\varphi, \sim \mathbf{p}_{j+1} \vdash \mathbf{A}$ by inductive hypothesis (using φ'')
- (.3) $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_j^\varphi \vdash \mathbf{p}_{j+1} \supset \mathbf{A}$ Deduction Theorem (1116): .1
- (.4) $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_j^\varphi \vdash \sim \mathbf{p}_{j+1} \supset \mathbf{A}$ Deduction Theorem (1116): .2
- (.5) $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_j^\varphi \vdash \mathbf{A}$ 1111: .3, .4

■

1205 Theorem. A wff is a theorem of \mathcal{P} if and only if it is a tautology.

Proof: By 1200 and 1204. ■

The *decision problem* for a logistic system is the problem of finding an effective procedure for determining of any wff A whether or not A is a theorem, and, if so, constructing a proof of A . Theorem 1205 provides a solution of the decision problem for \mathcal{P} , since one can effectively determine whether or not a wff is a tautology. Moreover, implicit in the proof of 1204 is a method for actually constructing a proof of any wff which is a tautology.

In showing that the theorems of \mathcal{P} are precisely the tautologies, we have established an important connection between the syntax and the semantics of \mathcal{P} . The set of theorems is characterized in two radically different ways, one syntactic (as the set of wffs which have proofs) and one semantic (as the set of tautologies). Note that the decision problem is stated in purely syntactic terms, but that a solution for it which relied solely on syntactic arguments about proofs might be very difficult. The introduction of semantic concepts greatly facilitates a solution to this syntactic problem.

EXERCISES

In exercises X1200–X1208, decide whether the wff is a tautology, a contradiction, or neither. If it is neither, give a falsifying assignment and a satisfying assignment. Use efficient methods.

$$\mathbf{X1200}. [\sim [p \wedge \sim q] \wedge [\sim r \vee s]] \supset [[s \supset q] \wedge \sim p] \supset \sim r$$

$$\mathbf{X1201}. [r \vee \neg q \wedge \sim \neg p \supset s] \supset \neg \sim [s \supset r] \supset \neg q \supset p$$

$$\mathbf{X1202}. p \equiv [p \vee p]$$

$$\mathbf{X1203}. [p \supset q] \equiv \neg [\sim p \vee q] \wedge \sim \neg p \wedge \sim q$$

$$\mathbf{X1204}. [p \supset \neg q \vee r] \supset \sim \neg [q \wedge \neg p \vee r] \supset \neg \sim r \supset \sim p$$

$$\mathbf{X1205}. \sim [p \vee q] \equiv \neg \sim p \wedge \sim q$$

$$\mathbf{X1206}. p \supset q \supset \sim \neg q \supset \sim p$$

$$\mathbf{X1207}. p \supset q \supset \neg q \supset [p \vee r] \supset \neg [p \vee r] \equiv q$$

$$\mathbf{X1208}. p \wedge [q \vee r] \equiv \neg [p \wedge q] \vee [p \wedge r]$$

X1209. If C and D are wffs, we say that D is obtained from C by *identifying*

certain propositional variables if \mathbf{D} is of the form $S_{q_1 \dots q_n}^{p_1 \dots p_n} C$, where p_1, \dots, p_n are distinct propositional variables of C and q_1, \dots, q_n are propositional variables of C which are distinct from p_1, \dots, p_n . Prove that if C is a tautology, and \mathbf{D} is obtained from C by identifying certain propositional variables, then \mathbf{D} is a tautology.

Justify your answers to the next two exercises with very rigorous proofs, using induction whenever the opportunity arises.

X1210. Does \mathcal{P} have any theorems in which there are no occurrences of disjunction?

X1211. Does \mathcal{P} have any theorems in which there are no occurrences of negation?

X1212. Let \mathcal{P}' be the system obtained from \mathcal{P} by adding the single wff $[p \supset q]$ to the axioms of \mathcal{P} . Show that \mathcal{P}' is consistent.

X1213. Suppose we add to \mathcal{P} the axiom schema $[\mathbf{A} \supset \mathbf{B}]$. Is the resulting system consistent?

X1214. Let \mathcal{P}_0 be the logistic system which has the same wffs and rule of inference as \mathcal{P} , and the following axiom schemas:

- (1) $[\mathbf{A} \wedge \sim \mathbf{A}] \supset \mathbf{B}$
- (2) $[\mathbf{A} \supset \mathbf{B}] \supset \bullet \mathbf{B} \supset \mathbf{A}$

Is \mathcal{P}_0 consistent?

X1215. Discuss the consistency and completeness of the following logistic system:

The wffs are those of \mathcal{P} . The axioms are all wffs of the form $[\mathbf{A} \supset S_{\mathbf{B}}^{\mathbf{p}} \mathbf{A}]$, where \mathbf{A} and \mathbf{B} are wffs and \mathbf{p} is any propositional variable. The sole rule of inference is Modus Ponens.

X1216. Let \mathcal{R} be a system of propositional calculus with a single binary connective $*$. Thus the wffs of \mathcal{R} are defined inductively as follows:

- (a) Every propositional variable is a wff.
- (b) If \mathbf{A} and \mathbf{B} are wffs, so is $[\mathbf{A} * \mathbf{B}]$.

The axioms of \mathcal{R} are all wffs of the form $[\mathbf{A} * \mathbf{A}]$. The sole rule of inference is: from \mathbf{A} and $[\mathbf{B} * \mathbf{A}]$ to infer \mathbf{B} . Solve the decision problem for \mathcal{R} , and prove that your solution is correct.

X1217. Let \mathcal{Q} be the system of propositional calculus whose sole primitive connectives are \sim (negation) and \wedge (conjunction), and which has the following rule of inference and axiom schemata:

Rule of Inference. From \mathbf{A} and $\sim \mathbf{A} \wedge \mathbf{B}$ to infer \mathbf{B} .

Axiom Schemata.

- (1) $\sim [\mathbf{A} \wedge \mathbf{A}] \wedge \mathbf{A}$
- (2) $\sim \mathbf{A} \wedge \neg \mathbf{B} \wedge \mathbf{A}$
- (3) $\sim [\sim \mathbf{A} \wedge \mathbf{B}] \wedge \neg \sim [\mathbf{C} \wedge \mathbf{A}] \wedge \neg \mathbf{B} \wedge \mathbf{C}$

Prove that a wff of \mathcal{Q} is a theorem of \mathcal{Q} if and only if it is a contradiction.

X1218. If \mathcal{Q} is any absolutely consistent system of propositional calculus, are the theorems of \mathcal{Q} necessarily all tautologies? Give a complete justification for your answer.

X1219. Let \mathcal{P}_1 be a logistic system with the same primitive symbols, wffs, abbreviations and rule of inference as \mathcal{P} . The two axiom schemata for \mathcal{P}_1 are:

$$\begin{aligned} & [[\mathbf{A} \supset \mathbf{B}] \supset \mathbf{A}] \supset \mathbf{A} \\ & [\mathbf{A} \vee [\mathbf{B} \wedge \mathbf{C}]] \supset \neg \mathbf{C} \vee \mathbf{A} \end{aligned}$$

Is every theorem of \mathcal{P}_1 a theorem of \mathcal{P} ?

X1220. Let \mathcal{P}_2 be obtained from \mathcal{P} by adding to \mathcal{P} as axioms all wffs of the form $\sim \mathbf{p}$ (where \mathbf{p} is a propositional variable).

- (a) Is \mathcal{P}_2 consistent?
- (b) Solve the decision problem for \mathcal{P}_2 .

X1221. Prove that if \mathbf{A} is any wff of \mathcal{P} and $\varphi \mathbf{p} = \psi \mathbf{p}$ for each variable \mathbf{p} which occurs in \mathbf{A} , then $\mathcal{V}_\varphi \mathbf{A} = \mathcal{V}_\psi \mathbf{A}$.

X1222. For any wff \mathbf{A} (written in primitive notation without abbreviations) of \mathcal{P} , let \mathbf{A}' (which is called the *dual* of \mathbf{A}) be the wff whose abbreviation is obtained when the symbol \vee is replaced by \wedge everywhere in \mathbf{A} . (Thus if \mathbf{A} is $\sim p \vee q$, then \mathbf{A}' is $\sim p \wedge q$, i.e., $\sim [\sim \sim p \vee \sim q]$.) You may use abbreviations freely in your work on this problem.

- (a) Give a definition of \mathbf{A}' by induction on the construction of \mathbf{A} .

If φ is any assignment of truth values to propositional variables, let φ' be the “opposite” assignment, i.e., $\varphi' \mathbf{p} = \mathbf{F}$ if $\varphi \mathbf{p} = \mathbf{T}$, and $\varphi' \mathbf{p} = \mathbf{T}$ if $\varphi \mathbf{p} = \mathbf{F}$.

By using \sim in the meta-language to denote the negation operation on truth values, we may concisely define φ' thus: $\varphi' p = \sim \varphi p$ for every propositional variable p .

- (b) Compute $\mathcal{V}_{\varphi'} \sim \mathbf{A}$ and $\mathcal{V}_\varphi \mathbf{A}$ when \mathbf{A} is $[\sim [p \vee \sim q] \vee p]$ and $\varphi p = F$ and $\varphi q = T$.
- (c) Give a rigorous proof that if \mathbf{A} is any wff of \mathcal{P} and φ is any assignment, then $\mathcal{V}_{\varphi'} \sim \mathbf{A}' = \mathcal{V}_\varphi \mathbf{A}$.

X1223. Let \mathcal{K} be the system which has the same wffs and rule of inference as \mathcal{P} , and the following axiom schemata:

1. $\mathbf{A} \supset \mathbf{A}$.
2. $\mathbf{A} \supset \mathbf{B} \supset \dots \sim \mathbf{B} \supset \sim \mathbf{A}$

Is the wff $[p \supset q] \supset \dots \sim p \supset \sim q$ a theorem of \mathcal{K} ?

X1224. Let \mathcal{P}^* be the system obtained from \mathcal{P} by adding the single wff $[p \supset \sim p]$ (which we shall call Axiom 4) to the axioms of \mathcal{P} . (In this problem $\vdash A$ means A is provable in \mathcal{P}^* .) Consider the following arguments:

Argument A: \mathcal{P}^* is consistent with respect to negation. For it is easy to see by induction on proofs that every theorem of \mathcal{P}^* is satisfied by the particular assignment φ which gives the value falsehood to every propositional variable. If $\vdash \mathbf{A}$ and $\vdash \sim \mathbf{A}$, then $\mathcal{V}_\varphi \mathbf{A} = T$ and $\mathcal{V}_\varphi \sim \mathbf{A} = T$, so $\mathcal{V}_\varphi \mathbf{A} = F$, which is a contradiction. Thus, there can be no such wff \mathbf{A} , and \mathcal{P}^* is consistent with respect to negation.

Argument B: \mathcal{P}^* is inconsistent with respect to negation.

- | | |
|--|------------------------------|
| $(1) \vdash p \supset \sim p$
$(2) \vdash [p \supset \sim p] \supset \sim [p \supset \sim p]$
$(3) \vdash \sim [p \supset \sim p]$ | Axiom 4
Sub: 1
MP: 1,2 |
|--|------------------------------|

Thus, there is a wff \mathbf{A} such that (1) $\vdash \mathbf{A}$ and (3) $\vdash \sim \mathbf{A}$, so \mathcal{P}^* is inconsistent with respect to negation.

Find a flaw in one of these arguments, and explain exactly what the problem is.

§13. Independence

An axiom, axiom schema, or rule of inference of a logistic system is said to be *independent* iff the set of theorems is diminished when it is deleted from the system. More precisely, an axiom of a logistic system is *independent* iff it cannot be derived from the other axioms by the rules of inference. An axiom schema is *independent* iff not every instance of it is a theorem of the system obtained by deleting from the set of axioms all instances of the schema in question. A rule of inference is *independent* if there is a theorem which cannot be derived without using that rule of inference.

We remark that in order to show the independence of an axiom, it suffices to find some property which that axiom does not have but which all the other axioms of the system do have, and which is preserved by the rules of inference. For in this case all theorems derived without using the axiom in question must have the property, and the axiom in question cannot be one of these theorems. It often happens that an appropriate property for this purpose is that of being valid under some particular (unsound) interpretation of the system.

1300 Theorem. The axiom schemata and the rule of inference of \mathcal{P} are all independent.

Proof: It is trivial to see the Modus Ponens is independent, since it is the sole rule of inference. Simply note that $p \vee \sim p$ is a theorem (1103), but not an axiom.

To prove the independence of the axiom schemata, we use three truth values, 0, 1, and 2. In each case we *designate* certain truth values, and call a wff *special* (under the given interpretation) iff its value (under that interpretation) is designated for all assignments of values from the set $\{0, 1, 2\}$ to its variables.

For Axiom Schema 1, we use the interpretation of \vee and \sim expressed in Figure 1.10. (Note that here $p \vee q = p \times q \bmod 4$.) Here 0 is designated. It is easy to check that all instances of Axiom Schemata 2 and 3 are special. (To simplify the analysis, use the fact that if \mathbf{A} or \mathbf{B} is 0, then $[\mathbf{A} \vee \mathbf{B}]$ is 0 also.) Also, if \mathbf{A} is 0 and $\sim \mathbf{A} \vee \mathbf{B}$ is 0, then \mathbf{B} must be 0, so Modus Ponens preserves specialness. However $\sim [p \vee p] \vee p$ is 2 when p is 2, so Axiom Schema 1 is independent.

To prove the independence of Axiom Schema 2, we let 0 and 1 be designated and use the interpretation of \vee and \sim in Figure 1.11. It is easily

\vee	0	1	2	\sim
0	0	0	0	1
1	0	1	2	0
2	0	2	0	2

Figure 1.10:
Table for Axiom 1

\vee	0	1	2	\sim
0	0	0	0	2
1	0	1	2	1
2	0	2	2	0

Figure 1.11:
Table for Axiom 2

\vee	0	1	2	\sim
0	0	0	0	1
1	0	1	1	2
2	0	0	2	0

Figure 1.12:
Table for Axiom 3

checked that all instances of Axiom Schemata 1 and 3 are special. If \mathbf{A} is 0 or 1 then $\sim \mathbf{A}$ is 2 or 1, so in order that $[\sim \mathbf{A} \vee \mathbf{B}]$ be 0 or 1 \mathbf{B} must be 0 or 1, so Modus Ponens preserves specialness. However, $\sim p \vee \blacksquare q \vee p$ is 2 when p is 1 and q is 2, so Axiom Schema 2 is independent.

To prove the independence of Axiom Schema 3, we use Figure 1.12⁷, with 0 designated. All instances of Axiom Schemata 1 and 2 are special, and if \mathbf{A} is 0 and $\sim \mathbf{A} \vee \mathbf{B}$ is 0, then \mathbf{B} must be 0, so Modus Ponens preserves specialness. However $\mathcal{V}_\varphi [\sim [\sim p \vee q] \vee \blacksquare \sim [r \vee p] \vee \blacksquare q \vee r] = 1$ when $\varphi p = 1 = \varphi q$ and $\varphi r = 2$, so this instance of Axiom Schema 3 is not derivable from Axioms 1 and 2. ■

EXERCISES

X1300. Let \mathcal{P}' be the logistic system with the same wffs as \mathcal{P} , with Modus Ponens and Substitution (i.e., 1101 for the case where \mathcal{H} is empty) as primitive rules of inference, and with the following three axioms:

- (1) $p \vee p \supset p$
 - (2) $p \supset \blacksquare q \vee p$
 - (3) $p \supset q \supset \blacksquare r \vee p \supset \blacksquare q \vee r$
- (a) Prove that \mathcal{P}' has exactly the same theorems as \mathcal{P} .
(b) Prove the independence of the rules of inference of \mathcal{P}' .

X1301. Using the methods of the proofs above, show that if any of the three axiom schemata of \mathcal{P} were replaced by the schema $\mathbf{A} \vee \sim \mathbf{A}$ or by $\sim \mathbf{A} \vee \mathbf{A}$, the resulting system would not be complete (i.e., would not have all tautologies as theorems).

⁷This table was found by Yuh-yun Marjorie Hsu.

X1302. Let \mathcal{P}^* be the system of propositional calculus with primitive connectives \sim and \supset , with Modus Ponens (from \mathbf{A} and $[\mathbf{A} \supset \mathbf{B}]$ to infer \mathbf{B}) as sole rule of inference, and the following axiom schemata:

$$\begin{aligned} & \mathbf{A} \supset \mathbf{B} \supset \mathbf{A} \\ & \mathbf{A} \supset [\mathbf{B} \supset \mathbf{C}] \supset \mathbf{A} \supset \mathbf{B} \supset \mathbf{A} \supset \mathbf{C} \\ & \sim \sim \mathbf{A} \supset \mathbf{A}. \end{aligned}$$

Prove that $p \supset \sim \sim p$ is not derivable in \mathcal{P}^* . (*Hint:* consider interpretations with two truth values in which \supset has the usual meaning, but \sim denotes a truth function other than negation.)

§14. Propositional Connectives

We now leave the system \mathcal{P} and turn to considerations which are relevant to a variety of formulations of propositional calculus. The reader who wishes to see some other formulations of propositional calculus may look in [Church, 1956], §§10, 25, 26, 29, and references cited therein.

An *n-ary truth-function* is a function from *n*-tuples of truth values to truth values. As a special case we regard truth values as 0-ary truth functions. A *propositional connective* is a symbol of a formalized language which, in the intended interpretation, denotes a truth function. Thus \sim , the negation connective, denotes the negation function, which maps truth to falsehood and falsehood to truth. Figure 1.13 lists all the truth functions of two arguments, together with connectives which denote them.

Various other names and notations are sometimes used for these connectives. Peirce's dagger is sometimes called "nor", and Sheffer's stroke is sometimes called "nand". It is common to use \Rightarrow for implication and \Leftrightarrow for equivalence, but we prefer to reserve these for informal use in our meta-language.

By a *wff of propositional calculus* we mean a wff of a formulation of propositional calculus which has among its primitive symbols all propositional connectives. Semantic notions such as tautology, $\mathcal{V}_\phi \mathbf{A}$, etc., are defined for arbitrary wffs of propositional calculus in a manner analogous to that in §12. Some basic tautologies are summarized at the end of this section.

An operator (connective or function) is called *unary* or *monadic* if it takes one argument (like \sim or cosine), *binary* if it takes two arguments (like \supset or $+$), and *n-ary* if it takes *n* arguments. It is customary to write

	p	T	T	F	F
	q	T	F	T	F
Falsehood	f	F	F	F	F
Non-disjunction, Peirce's dagger	$p \downarrow q$	F	F	F	T
Converse non-implication	$p \not\subset q$	F	F	T	F
Negation		$\sim p$	F	F	T
Non-implication		$p \not\rightarrow q$	F	T	F
Negation		$\sim q$	F	T	F
(Material) non-equivalence, exclusive disjunction	$p \neq q$	F	T	T	F
Non-conjunction, Sheffer's stroke	$p \mid q$	F	T	T	T
Conjunction	$p \wedge q$	T	F	F	F
(Material) equivalence	$p \equiv q$	T	F	F	T
	q	T	F	T	F
(Material) implication	$p \supset q$	T	F	T	T
	p	T	T	F	F
Converse implication	$p \subset q$	T	T	F	T
(Inclusive) disjunction, alternation	$p \vee q$	T	T	T	F
Truth	t	T	T	T	T

Figure 1.13: Binary truth functions.

n -ary operators in *prefix* position *in front of* their arguments (as in $\sim p$ and $\cos \pi$ and $f(x_1, \dots, x_n)$) except when $n = 2$. Binary operators are customarily written in *infix* position *between* their arguments (as in $[p \supset q]$ and $3 + 5$). However, in theoretical discussions it is sometimes convenient to pretend that all operators are written in prefix position, and write $* \mathbf{B}_1 \mathbf{B}_2$ as a shorthand for $[\mathbf{B}_1 * \mathbf{B}_2]$. Thus, every wff of propositional calculus may be regarded as having the form \mathbf{p} or $* \mathbf{B}_1 \dots \mathbf{B}_n$, where $*$ is an n -ary propositional connective. Notice that when one uses prefix notation and knows the number of arguments of each operator, one does not really need brackets. In prefix notation $\supset\supset pqr$ means $[[p \supset q] \supset r]$, while $\supset p \supset qr$ means $[p \supset [q \supset r]]$.⁸

Theorems 1401 and 1402 below present facts involving substitutions and tautologies which seem so obvious that they are sometimes used without justification. Theorem 1400 serves as a lemma for 1401.

⁸Prefix notation for binary propositional connectives is sometimes referred to as *Polish* notation, since it was actually used by certain Polish logicians.

DEFINITION. A *proper substitution* is a substitution of the form $S_{C_1 \dots C_n}^{P_1 \dots P_n}$.

DEFINITION. If θ is a proper substitution, and φ is any assignment, the assignment $\varphi \circ \theta$ is defined as follows:

$$(\varphi \circ \theta) p = \mathcal{V}_\varphi \theta p \text{ for every propositional variable } p.$$

EXAMPLE: If $\varphi q = T$ and $\varphi r = F$ and $\theta = S_{[q \supset r]}^p$, then $(\varphi \circ \theta)p = \mathcal{V}_\varphi(\theta p) = \mathcal{V}_\varphi[q \supset r] = F$.

1400 Substitution-Value Theorem. Let φ be any assignment and θ be any proper substitution. If A is any wff of propositional calculus, then $\mathcal{V}_\varphi \theta A = \mathcal{V}_{\varphi \circ \theta} A$.

EXAMPLE: Let φ and θ be as in the previous example. Note that $\theta q = q$ and $\theta r = r$ so $(\varphi \circ \theta)q = T$ and $(\varphi \circ \theta)r = F$. Let W1400 be $[[p \supset r] \wedge \neg q \vee p]$. Then $\mathcal{V}_\varphi \theta W1400 = \mathcal{V}_\varphi [[(q \supset r) \supset r] \wedge \neg q \vee \neg q \supset r] = T$

$$\begin{array}{ccccccc} T & F & F & T & F & T & T \\ F & T & F & T & F & T & F \end{array}$$

$$\text{and } \mathcal{V}_{\varphi \circ \theta} W1400 = \mathcal{V}_{\varphi \circ \theta} [[p \supset r] \wedge \neg q \vee p] = T$$

Proof of 1400 by induction on the construction of A:

Case: A is a propositional variable p . $\mathcal{V}_\varphi \theta A = \mathcal{V}_\varphi \theta p = (\varphi \circ \theta)p$ (by the definition of $\varphi \circ \theta$) $= \mathcal{V}_{\varphi \circ \theta} A$

Case: A has the form $*B_1 \dots B_n$, where $*$ is an n -ary propositional connective denoting the truth function \oplus , with $n \geq 1$.

$$\begin{aligned} \mathcal{V}_\varphi \theta A &= \mathcal{V}_\varphi * (\theta B_1) \dots (\theta B_n) && \text{(by the definition of a substitution)} \\ &= \oplus (\mathcal{V}_\varphi \theta B_1) \dots (\mathcal{V}_\varphi \theta B_n) && \text{(by the definition of } \mathcal{V}_\varphi \text{)} \\ &= \oplus (\mathcal{V}_{\varphi \circ \theta} B_1) \dots (\mathcal{V}_{\varphi \circ \theta} B_n) && \text{(by inductive hypothesis)} \\ &= \mathcal{V}_{\varphi \circ \theta} * B_1 \dots B_n && \text{(by the definition of } \mathcal{V}_{\varphi \circ \theta} \text{)} \\ &= \mathcal{V}_{\varphi \circ \theta} A \end{aligned} \quad \blacksquare$$

1401 Corollary If $\models A$, and θ is any proper substitution, then $\models \theta A$.

Proof: Let φ be any assignment. $\mathcal{V}_\varphi \theta A = \mathcal{V}_{\varphi \circ \theta} A$ (by 1400) $= T$ (since $\models A$). Thus every assignment satisfies θA , so $\models \theta A$. \blacksquare

p	q	$p \supset q$	$\supset q$
T	T	T	T
F	F	F	T
F	T	T	T
F	F	T	F

Figure 1.14: \vee expressed in terms of \supset

1402 Substitutivity of Equivalence. If A , B , C , and D are wffs of propositional calculus, and D is obtained from C by replacing zero or more occurrences of A in C by occurrences of B , and $\models [A \equiv B]$, then $\models [C \equiv D]$.

Proof: By induction on the number of occurrences of propositional connectives in C . Consider the possible forms that C can have, and also the special cases where D is C , and where C is A and D is B .

NOTATION. If p_1, \dots, p_n are distinct propositional variables including all variables in A , we let $[\lambda p_1 \dots \lambda p_n A]$ denote that truth function whose value on any n -tuple (x_1, \dots, x_n) of truth values is $\nu_\varphi A$ when $\varphi p_i = x_i$ for $1 \leq i \leq n$.

EXAMPLE: It can be seen with the aid of Figure 1.14 that $[\lambda p \lambda q \cdot p \supset q \supset q]$ is the truth function h such that $h(T, T) = T$, $h(T, F) = T$, $h(F, T) = T$, and $h(F, F) = F$. This is commonly known as the disjunction function, and denoted by \vee .

Clearly, $[\lambda p_1 \dots \lambda p_n A]$ and $[\lambda p_1 \dots \lambda p_n B]$ denote the same truth function iff $[A \equiv B]$ is a tautology. (Exercise X1418.) For example, $\models [p \vee q] \equiv [p \supset q \supset q]$, so $[\lambda p \lambda q \cdot p \vee q] = [\lambda p \lambda q \cdot p \supset q \supset q]$.

NOTATION. We shall use $\bigwedge_{i=1}^m A_i$ as a notation for $[A_1 \wedge \dots \wedge A_m]$, and $\bigvee_{i=1}^m A_i$ as a notation for $[A_1 \vee \dots \vee A_m]$. For $m \geq 1$ these notations can be defined inductively: $\bigwedge_{i=1}^1 A_i = A_1$ and $\bigwedge_{i=1}^{m+1} A_i = \bigwedge_{i=1}^m A_i \wedge A_{m+1}$; similarly, $\bigvee_{i=1}^1 A_i = A_1$ and $\bigvee_{i=1}^{m+1} A_i = \bigvee_{i=1}^m A_i \vee A_{m+1}$.

Sometimes we need more general notations for a disjunction or conjunction of a finite set of wffs. We discuss general notations for disjunctions; similar notations will be used for conjunctions. If P is a finite index set $\{i_1, \dots, i_n\}$ or a finite sequence $< i_1, \dots, i_n >$, and A_k is a wff for each $k \in P$, then $\bigvee_{k \in P} A_k$ stands for $[A_{i_1} \vee \dots \vee A_{i_n}]$. If P is a finite set or sequence of wffs (so that $A_k = k$ in the notation above), instead of

$\bigvee_{k \in P} k$ we may simply write $\bigvee P$ for the disjunction of the wffs in P . Of course, in case P is a finite set, these notations introduce mild ambiguity, since (for example) $\{1, 3, 2, 1\} = \{3, 1, 2\}$, but disjunction and conjunction are associative and commutative, so $\bigvee P$ with one representation of P will be equivalent to $\bigvee P$ with any other representation of P .

In certain contexts it is convenient to form conjunctions and disjunctions when the index set P is empty or $m = 0$, in the notation above. An empty conjunction is always true, and may be symbolized by t , while an empty disjunction is always false, and may be symbolized by f or \square . These conventions are justified by the fact that adding on an empty conjunction or disjunction to a conjunction or disjunction does not alter its truth, since $A \wedge B \wedge t \equiv A \wedge B$ and $A \vee B \vee f \equiv A \vee B$.

The following simple but very useful tautologies are known as DeMorgan's laws: $[\sim [A_1 \wedge A_2] \equiv \neg \sim A_1 \vee \sim A_2]$ and $[\sim [A_1 \vee A_2] \equiv \neg \sim A_1 \wedge \neg \sim A_2]$. Using the notation above, they can be stated in more general form as $[\sim \bigwedge_{i=1}^m A_i \equiv \bigvee_{i=1}^m \sim A_i]$ and $[\sim \bigvee_{i=1}^m A_i \equiv \bigwedge_{i=1}^m \sim A_i]$. Note that the generalized DeMorgan laws are correct even when $m = 0$ or $m = 1$.

DEFINITIONS.

- (1) A *literal* is a wff of the form p or of the form $\sim p$.
- (2) A wff D is in *disjunctive normal form (dnf)* iff it is a disjunction of conjunctions of literals. Thus it is of the form $\bigvee_{i=1}^m \bigwedge_{j=1}^{n_i} P_{ij}$, where the P_{ij} are literals. The conjunctions of literals $\bigwedge_{j=1}^{n_i} P_{ij}$ which are the scopes of the disjunctions of D are called the *disjuncts* of D . As a special case we admit the possibility that m or any n_i may be 1. Similarly, a wff is in *conjunctive normal form (cnf)* iff it is a conjunction of disjunctions of literals, and the disjunctions of literals which are the scopes of its conjunctions are called its *conjuncts*.
- (3) A wff is in *full disjunctive normal form* iff it has the form $\bigvee_{i=1}^m \bigwedge_{j=1}^n P_{ij}$, where for each i , P_{ij} is p_j or $\sim p_j$, and p_1, \dots, p_n are distinct propositional variables. Full conjunctive normal form is defined similarly.
- (4) A wff A is a *disjunctive normal form for a wff B* iff A is in disjunctive normal form and is equivalent to B . Similarly, A is a *conjunctive normal form for a wff B* iff A is in conjunctive normal form and is equivalent to B .

EXAMPLES: $p \vee [q \wedge \sim p] \vee [\sim q \wedge r \wedge \sim r]$ is in disjunctive normal form, and $\sim p \wedge [\sim q \vee p] \wedge [q \vee \sim r \vee r]$ is in conjunctive normal form.

			$\sim \bullet [q \equiv r] \supset p$	Conjunctions of literals for disjunctive normal form
T	T	T	F	
T	T	F	T	$q \wedge r \wedge \sim p$
T	F	T	F	
T	F	F	F	
F	T	T	F	
F	T	F	F	
F	F	T	F	
F	F	F	T	$\sim q \wedge \sim r \wedge \sim p$

Figure 1.15: Conjunctions of literals for dnf of $\sim \bullet [q \equiv r] \supset p$

$[p \wedge q \wedge \sim r] \vee [p \wedge q \wedge r] \vee [\sim p \wedge \sim q \wedge r]$ is in full disjunctive normal form.

EXAMPLE: $p \equiv q$ has disjunctive normal form $[p \wedge q] \vee [\sim p \wedge \sim q]$ and conjunctive normal form $[\sim p \vee q] \wedge [p \vee \sim q]$.

Next we shall show that every wff of propositional calculus has disjunctive and conjunctive normal forms. The basic idea in the proof below can be illustrated by using the truth table in Figure 1.15 to find a disjunctive normal form for $\sim \bullet [q \equiv r] \supset p$. We find the lines of the truth table (which correspond to assignments of values to variables) which make the wff true, and represent these as conjunctions of literals. The disjunction of these conjunctions of literals is a dnf for the wff. In our example, examination reveals that the wff is true iff q and r are true, and p is false, or q , r , and p are all false, so the disjunctive normal form is

$$[q \wedge r \wedge \sim p] \vee [\sim q \wedge \sim r \wedge \sim p].$$

1403 Theorem. Let h be any truth function with n arguments, where $n \geq 1$, and let p_1, \dots, p_n be distinct propositional variables. Then there is a wff \mathbf{A} in disjunctive normal form such that $h = [\lambda p_1, \dots, \lambda p_n \mathbf{A}]$.

Proof: If the value of h is F for all arguments, let \mathbf{A} be $p_1 \wedge \sim p_1$.

Otherwise, for each assignment φ let $\mathbf{A}_\varphi = p_1^\varphi \wedge \dots \wedge p_n^\varphi$, where

$$p_i^\varphi = p_i \text{ if } \varphi p_i = T$$

$$= \sim p_i \text{ if } \varphi p_i = F.$$

Thus $\mathcal{V}_\varphi \mathbf{A}_\varphi = \top$, and if ψ is any assignment such that $\mathcal{V}_\psi \mathbf{A}_\varphi = \top$, then $\psi \mathbf{p}_i = \varphi \mathbf{p}_i$ for each $i \leq n$.

Let $\mathbf{A} = \bigvee_{\varphi \in \{\varphi \mid h(\varphi \mathbf{p}_1, \dots, \varphi \mathbf{p}_n) = \top\}} \mathbf{A}_\varphi$; thus \mathbf{A} is the disjunction of those wffs \mathbf{A}_φ such that $h(\varphi \mathbf{p}_1, \dots, \varphi \mathbf{p}_n) = \top$. Note that for any assignment ψ , $\mathcal{V}_\psi \mathbf{A} = h(\psi \mathbf{p}_1, \dots, \psi \mathbf{p}_n)$, since $\mathcal{V}_\psi \mathbf{A} = \top$ iff $h(\psi \mathbf{p}_1, \dots, \psi \mathbf{p}_n) = \top$, and there are only two truth values.

To show that $[\lambda \mathbf{p}_1 \dots \lambda \mathbf{p}_n \mathbf{A}] = h$, we let (x_1, \dots, x_n) be any n -tuple of truth values, and show that $[\lambda \mathbf{p}_1 \dots \lambda \mathbf{p}_n \mathbf{A}](x_1 \dots x_n) = h(x_1, \dots, x_n)$. Define ψ so that $\psi \mathbf{p}_i = x_i$ for $1 \leq i \leq n$. Then $[\lambda \mathbf{p}_1 \dots \lambda \mathbf{p}_n \mathbf{A}](x_1 \dots x_n) = \mathcal{V}_\psi \mathbf{A} = h(\psi \mathbf{p}_1, \dots, \psi \mathbf{p}_n) = h(x_1, \dots, x_n)$. ■

EXAMPLE: Let $h(x_1, x_2) = \top$ iff $x_1 = x_2$. Let $\varphi_1 \mathbf{p}_1 = \varphi_1 \mathbf{p}_2 = \top$ and $\varphi_2 \mathbf{p}_1 = \varphi_2 \mathbf{p}_2 = \perp$. Then $\mathbf{A} = \mathbf{A}_{\varphi_1} \vee \mathbf{A}_{\varphi_2} = [\mathbf{p}_1 \wedge \mathbf{p}_2] \vee [\sim \mathbf{p}_1 \wedge \sim \mathbf{p}_2]$.

1404 Corollary. Every wff of propositional calculus has a disjunctive normal form.

Proof: Let \mathbf{B} be an arbitrary wff, let $\mathbf{p}_1, \dots, \mathbf{p}_n$ be the propositional variables in \mathbf{B} , and let $h = [\lambda \mathbf{p}_1 \dots \lambda \mathbf{p}_n \mathbf{B}]$. By 1403 there is a wff \mathbf{A} in disjunctive normal form such that $h = [\lambda \mathbf{p}_1 \dots \lambda \mathbf{p}_n \mathbf{A}]$, so $\vdash \mathbf{B} \equiv \mathbf{A}$.

REMARK. The following tautologies, known as distributive laws, are often useful for putting wffs into disjunctive or conjunctive normal form:

$$[\mathbf{A} \vee \mathbf{B}] \wedge \mathbf{C} \equiv [\mathbf{A} \wedge \mathbf{C}] \vee [\mathbf{B} \wedge \mathbf{C}]$$

$$[\mathbf{A} \wedge \mathbf{B}] \vee \mathbf{C} \equiv [\mathbf{A} \vee \mathbf{C}] \wedge [\mathbf{B} \vee \mathbf{C}].$$

EXAMPLE: Find a disjunctive normal form for $\sim \bullet [q \equiv r] \supset p$. We have already shown how to do this using the method implicit in the proofs of theorems 1403 and 1404. Another method is to transform the given wff into disjunctive normal form step by step, using simple tautologies and the Substitutivity of Equivalence, as follows:

$$\sim \bullet [q \equiv r] \supset p$$

$$[q \equiv r] \wedge \sim p$$

$$[[q \wedge r] \vee \sim q \wedge \sim r] \wedge \sim p$$

$$[q \wedge r \wedge \sim p] \vee \sim q \wedge \sim r \wedge \sim p.$$

Informally, we say that a set of propositional connectives is complete iff every truth function can be expressed in terms of those connectives. This can be stated more formally as follows:

DEFINITION. A set \mathcal{S} of propositional connectives is *complete* iff for each integer $n \geq 1$ and for each truth function h of n arguments, there is a wff \mathbf{A} in which only connectives of \mathcal{S} occur such that $h = [\lambda p_1 \dots \lambda p_n \mathbf{A}]$.

1405 Theorem. The following sets of propositional connectives are complete:

- | | |
|-------------------------|---------------------------|
| (a) $\{\vee, \sim\}$ | (e) $\{ \}$ |
| (b) $\{\wedge, \sim\}$ | (f) $\{\downarrow\}$ |
| (c) $\{\supset, \sim\}$ | (g) $\{\wedge, \neq, t\}$ |
| (d) $\{\supset, f\}$ | |

Proof: It follows directly from theorem 1403 that $\{\wedge, \vee, \sim\}$ is complete. However, since for any wffs \mathbf{C} and \mathbf{D} , $[\mathbf{C} \wedge \mathbf{D}]$ is equivalent to $\sim [\sim \mathbf{C} \vee \sim \mathbf{D}]$, any wff in which \sim , \vee , and \wedge are the only connectives is equivalent to some wff in which \sim and \vee are the only connectives. This proves (a). The remaining assertions follow by similar arguments using the following equivalences:

- (b) $[\mathbf{A} \vee \mathbf{B}] \equiv \sim \neg \mathbf{A} \wedge \neg \mathbf{B}$
- (c) $[\mathbf{A} \vee \mathbf{B}] \equiv \neg \neg \mathbf{A} \supset \mathbf{B}$
- (d) $\neg \mathbf{A} \equiv \neg \mathbf{A} \supset f$
- (e) $\neg \mathbf{A} \equiv [\mathbf{A} \mid \mathbf{A}]$ and $[\mathbf{A} \wedge \mathbf{B}] \equiv \neg [\mathbf{A} \mid \mathbf{B}]$
- (f) $\neg \mathbf{A} \equiv [\mathbf{A} \downarrow \mathbf{A}]$ and $[\mathbf{A} \vee \mathbf{B}] \equiv \neg [\mathbf{A} \downarrow \mathbf{B}]$.
- (g) $\neg \mathbf{A} \equiv \neg \mathbf{A} \neq t$

■

DEFINITION. A wff of propositional calculus is in *negation normal form (nnf)*, and is called a *negation normal formula* (which we also abbreviate as *nnf*), iff it contains no propositional connectives other than \wedge , \vee , and \sim , and the scope of each occurrence of \sim is a propositional variable.

It follows from 1404 that every wff of propositional calculus is equivalent to a wff in nnf. However, a negation normal form of a wff may be much shorter than any disjunctive normal form of that wff. A wff can be put into nnf very easily by expressing all its connectives in terms of \wedge , \vee , and \sim , and then applying DeMorgan's Laws and eliminating double negations until the scopes of all negations are propositional variables.

$$\left[\left[\begin{array}{c} p \\ q \end{array} \right] \vee \left[\begin{array}{c} \sim p \\ \sim q \end{array} \right] \right]$$

Figure 1.16: dnf of $p \equiv q$

$$\left[\begin{array}{c} \sim p \vee q \\ p \vee \sim q \end{array} \right]$$

Figure 1.17: cnf of $p \equiv q$

$$\left[\left[\begin{array}{c} p \vee q \\ \sim r \\ \sim q \vee s \end{array} \right] \vee \left[\begin{array}{c} \sim p \\ r \vee \sim s \end{array} \right] \right]$$

Figure 1.18: Vpform for W1401: $[(p \vee q) \wedge \sim r \wedge \sim q \vee s] \vee [\sim p \wedge r \vee \sim s]$

It is sometimes illuminating to display nnfs in a two-dimensional format called a *vpform* (vertical path form), in which disjunctions are written horizontally in the usual way, but conjunctions are written vertically, so that $[\mathbf{A} \wedge \mathbf{B}]$ is written as $\left[\begin{array}{c} \mathbf{A} \\ \mathbf{B} \end{array} \right]$.

Vpforms for the dnf and cnf of $p \equiv q$ are displayed in Figures 1.16 and 1.17. A vpform for $[(p \vee q) \wedge \sim r \wedge \sim q \vee s] \vee [\sim p \wedge r \vee \sim s]$, which we call W1401, is displayed in Figure 1.18.

It is useful to define the *vertical* and *horizontal paths* through a nnf \mathbf{A} as certain sequences of occurrences of literals in that nnf. We shall let $\mathcal{VP}(\mathbf{A})$ be the set of vertical paths through \mathbf{A} , and $\mathcal{HP}(\mathbf{A})$ be the set of horizontal paths through \mathbf{A} . Essentially, if one erases the left scope or the right scope of each occurrence of \vee (perhaps making different choices for different occurrences of \vee) in a nnf, the remaining literal occurrences form a vertical path. Similarly, if one erases the left scope or the right scope of each occurrence of \wedge in a nnf, the remaining literal occurrences form a horizontal path.

EXAMPLE: $\mathcal{VP}(W1401) = \{< p, \sim r, \sim q >, < p, \sim r, s >, < q, \sim r, \sim q >, < q, \sim r, s >, < \sim p, r >, < \sim p, \sim s >\}$.

$\mathcal{HP}(W1401) = \{< p, q, \sim p >, < p, q, r, \sim s >, < \sim r, \sim p >, < \sim r, r, \sim s >, < \sim q, s, \sim p >, < \sim q, s, r, \sim s >\}$.

The reader who traces these paths on Figure 1.18 will see that the ideas underlying vertical and horizontal paths are very simple and natural. We shall see that they are also very useful.

In order to prove theorems about vertical and horizontal paths, we need careful definitions of these concepts. Instead of directly defining what a

vertical (or horizontal) path is, we define the *set* of vertical (or horizontal) paths through a nnf. (This is a technique which is very often useful in mathematics.) If $P = \langle P_1, \dots, P_m \rangle$ and $Q = \langle Q_1, \dots, Q_n \rangle$ are paths, we shall use the notation PQ for the concatenation $\langle P_1, \dots, P_m, Q_1, \dots, Q_n \rangle$ of these paths.

DEFINITIONS. Let \mathbf{A} be a nnf. We define $\mathcal{VP}(\mathbf{A})$ and $\mathcal{HP}(\mathbf{A})$ by induction on the construction of \mathbf{A} as a negation normal formula:

- (1) If \mathbf{A} is a literal, $\mathcal{VP}(\mathbf{A}) = \{\langle \mathbf{A} \rangle\}$, i.e., the set containing the sequence of length 1 whose first and only member is the occurrence of \mathbf{A} in \mathbf{A} .
- (2) If \mathbf{A} has the form $[B \vee C]$, then $\mathcal{VP}(\mathbf{A}) = \mathcal{VP}(B) \cup \mathcal{VP}(C)$.
- (3) If \mathbf{A} has the form $[B \wedge C]$, then $\mathcal{VP}(\mathbf{A}) = \{PQ \mid P \in \mathcal{VP}(B) \text{ and } Q \in \mathcal{VP}(C)\}$.
- (1') If \mathbf{A} is a literal, $\mathcal{HP}(\mathbf{A}) = \{\langle \mathbf{A} \rangle\}$.
- (2') If \mathbf{A} has the form $[B \vee C]$, then $\mathcal{HP}(\mathbf{A}) = \{PQ \mid P \in \mathcal{HP}(B) \text{ and } Q \in \mathcal{HP}(C)\}$.
- (3') If \mathbf{A} has the form $[B \wedge C]$, $\mathcal{HP}(\mathbf{A}) = \mathcal{HP}(B) \cup \mathcal{HP}(C)$.

Associated with each sequence P of *occurrences* of literals in a wff is a sequence P' of *literals*. If a term of P is an *occurrence* of the literal L , then the corresponding term of P' is the literal L itself. We shall generally not distinguish carefully between sequences of literal occurrences and sequences of literals, use the same notation for P and P' , and rely on the reader to determine from the context which is intended.⁹ Thus, if P is a path through a nnf and φ is an assignment, we say that φ satisfies [falsifies] every literal in P iff φ satisfies [falsifies] every literal which has an occurrence in P .

1406 Theorem. Let \mathbf{A} be a wff in negation normal form, and let φ be an assignment. Then

- (a) φ satisfies \mathbf{A} iff there is a path $P \in \mathcal{VP}(\mathbf{A})$ such that φ satisfies every literal in P .
- (b) φ falsifies \mathbf{A} iff there is a path $P \in \mathcal{HP}(\mathbf{A})$ such that φ falsifies every literal in P .

Proof of 1406(a) by induction on the construction of \mathbf{A} :

⁹It might seem simpler to define vertical and horizontal paths as sequences of literals or as sets of literals, but this would not permit adequate descriptions of certain algorithms involving paths.

Case 1. \mathbf{A} is a literal. This case is trivial.

Case 2. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$. The following are equivalent:

- (i) $\mathcal{V}_\varphi \mathbf{A} = \top$.
- (ii) $\mathcal{V}_\varphi [\mathbf{B} \vee \mathbf{C}] = \top$.
- (iii) $\mathcal{V}_\varphi \mathbf{B} = \top$ or $\mathcal{V}_\varphi \mathbf{C} = \top$ (by the definition of \mathcal{V}_φ)
- (iv) there is a path $P \in \mathcal{VP}(\mathbf{B})$ such that φ satisfies every literal in P , or there is a path $P \in \mathcal{VP}(\mathbf{C})$ such that φ satisfies every literal in P (by the inductive hypotheses, which we can apply since \mathbf{B} and \mathbf{C} are in nnf).
- (v) there is a path $P \in \mathcal{VP}(\mathbf{B}) \cup \mathcal{VP}(\mathbf{C})$ such that φ satisfies every literal in P (by the definition of \cup).
- (vi) there is a path $P \in \mathcal{VP}(\mathbf{A})$ such that φ satisfies every literal in P (by the definition of $\mathcal{VP}([\mathbf{B} \vee \mathbf{C}])$).

Case 3. \mathbf{A} is $[\mathbf{B} \wedge \mathbf{C}]$. The following are equivalent:

- (i) $\mathcal{V}_\varphi \mathbf{A} = \top$.
- (ii) $\mathcal{V}_\varphi [\mathbf{B} \wedge \mathbf{C}] = \top$.
- (iii) $\mathcal{V}_\varphi \mathbf{B} = \top$ and $\mathcal{V}_\varphi \mathbf{C} = \top$ (by the definition of \mathcal{V}_φ)
- (iv) there is a path $P \in \mathcal{VP}(\mathbf{B})$ such that φ satisfies every literal in P , and there is a path $Q \in \mathcal{VP}(\mathbf{C})$ such that φ satisfies every literal in Q (by the inductive hypotheses).
- (v) there is a path $R \in \mathcal{VP}(\mathbf{A})$ such that φ satisfies every literal in R .

To prove that (iv) implies (v), note that if we have paths $P \in \mathcal{VP}(\mathbf{B})$ such that φ satisfies every literal in P and $Q \in \mathcal{VP}(\mathbf{C})$ such that φ satisfies every literal in Q , then their concatenation PQ is a path in $\mathcal{VP}([\mathbf{B} \wedge \mathbf{C}])$ such that φ satisfies every literal in PQ . The proof that (v) implies (iv) is also a direct consequence of the definition of $\mathcal{VP}([\mathbf{B} \wedge \mathbf{C}])$.

This completes the proof of (a). As might be expected, the proof of (b) mirrors the proof of (a).

Proof of 1406(b) by induction on the construction of \mathbf{A} :

Case 1. \mathbf{A} is a literal. This case is trivial.

Case 2. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$. The following are equivalent:

- (i) $\mathcal{V}_\varphi \mathbf{A} = \mathbb{F}$.

- (ii) $\mathcal{V}_\varphi [\mathbf{B} \vee \mathbf{C}] = \mathbf{F}$.
- (iii) $\mathcal{V}_\varphi \mathbf{B} = \mathbf{F}$ and $\mathcal{V}_\varphi \mathbf{C} = \mathbf{F}$ (by the definition of \mathcal{V}_φ).
- (iv) there is a path $P \in \mathcal{HP}(\mathbf{B})$ such that φ falsifies every literal in P , and there is a path $Q \in \mathcal{HP}(\mathbf{C})$ such that φ falsifies every literal in Q (by the inductive hypotheses).
- (v) there is a path $R \in \mathcal{HP}(\mathbf{A})$ such that φ falsifies every literal in R .

To prove that (iv) implies (v), note that if we have paths $P \in \mathcal{HP}(\mathbf{B})$ such that φ falsifies every literal in P and $Q \in \mathcal{HP}(\mathbf{C})$ such that φ falsifies every literal in Q , then their concatenation PQ is a path in $\mathcal{HP}([\mathbf{B} \vee \mathbf{C}])$ such that φ falsifies every literal in PQ . The proof that (v) implies (iv) is also a direct consequence of the definition of $\mathcal{HP}([\mathbf{B} \vee \mathbf{C}])$.

Case 3. \mathbf{A} is $[\mathbf{B} \wedge \mathbf{C}]$. The following are equivalent:

- (i) $\mathcal{V}_\varphi \mathbf{A} = \mathbf{F}$.
- (ii) $\mathcal{V}_\varphi [\mathbf{B} \wedge \mathbf{C}] = \mathbf{F}$.
- (iii) $\mathcal{V}_\varphi \mathbf{B} = \mathbf{F}$ or $\mathcal{V}_\varphi \mathbf{C} = \mathbf{F}$ (by the definition of \mathcal{V}_φ).
- (iv) there is a path $P \in \mathcal{HP}(\mathbf{B})$ such that φ falsifies every literal in P , or there is a path $P \in \mathcal{HP}(\mathbf{C})$ such that φ falsifies every literal in P (by the inductive hypotheses).
- (v) there is a path $P \in \mathcal{HP}(\mathbf{B}) \cup \mathcal{HP}(\mathbf{C})$ such that φ falsifies every literal in P .
- (vi) there is a path $P \in \mathcal{HP}(\mathbf{A})$ such that φ falsifies every literal in P (by the definition of $\mathcal{HP}([\mathbf{B} \wedge \mathbf{C}])$). ■

Next we shall see that we can construct a disjunctive normal form for a nnf simply by looking at its vpform, forming the conjunctions of the literals in each vertical path, and forming the disjunction of these conjunctions. A conjunctive normal form can be obtained by a similar procedure.

1407 Theorem. Let \mathbf{A} be a wff in negation normal form. Then

- (a) $\bigvee_{P \in \mathcal{VP}(\mathbf{A})} \bigwedge P$ is a disjunctive normal form of \mathbf{A} .
- (b) $\bigwedge_{P \in \mathcal{HP}(\mathbf{A})} \bigvee P$ is a conjunctive normal form of \mathbf{A} .

EXAMPLE:

A dnf of W1401 is $[[p \wedge \sim r \wedge \sim q] \vee [p \wedge \sim r \wedge s] \vee [q \wedge \sim r \wedge \sim q] \vee [q \wedge \sim r \wedge s] \vee [\sim p \wedge r] \vee [\sim p \wedge \sim s]]$.

A cnf of W1401 is $[[p \vee q \vee \sim p] \wedge [p \vee q \vee r \vee \sim s] \wedge [\sim r \vee \sim p] \wedge [\sim r \vee r \vee \sim s] \wedge [\sim q \vee s \vee \sim p] \wedge [\sim q \vee s \vee r \vee \sim s]]$.

The reader should note how the paths in Figure 1.18 occur in these wffs.

Proof of 1407(a): Clearly $\bigvee_{P \in \mathcal{VP}(\mathbf{A})} \bigwedge P$ is in disjunctive normal form. To show that it is equivalent to \mathbf{A} , let φ be any assignment. The following are equivalent:

- (i) φ satisfies \mathbf{A} .
- (ii) There is a path $P \in \mathcal{VP}(\mathbf{A})$ such that φ satisfies every literal in P (by 1406(a)).
- (iii) There is a path $P \in \mathcal{VP}(\mathbf{A})$ such that φ satisfies $\bigwedge P$.
- (iv) φ satisfies $\bigvee_{P \in \mathcal{VP}(\mathbf{A})} \bigwedge P$.

Proof of 1407(b): Clearly $\bigwedge_{P \in \mathcal{HP}(\mathbf{A})} \bigvee P$ is in conjunctive normal form. To show that it is equivalent to \mathbf{A} , let φ be any assignment. The following are equivalent:

- (i) φ falsifies \mathbf{A} .
- (ii) There is a path $P \in \mathcal{HP}(\mathbf{A})$ such that φ falsifies every literal in P (by 1406(b)).
- (iii) There is a path $P \in \mathcal{HP}(\mathbf{A})$ such that φ falsifies $\bigvee P$.
- (iv) φ falsifies $\bigwedge_{P \in \mathcal{HP}(\mathbf{A})} \bigvee P$. ■

EXAMPLE: Find a disjunctive normal form for $[[p \equiv q] \equiv \sim r]$.

The following are equivalent:

$$\begin{aligned}
 & [[p \equiv q] \equiv \sim r] \\
 & \left[\left[\begin{array}{c} p \equiv q \\ \sim r \end{array} \right] \vee \left[\begin{array}{c} p \not\equiv q \\ r \end{array} \right] \right] \\
 & \left[\left[\begin{array}{c} p \equiv q \\ \sim r \end{array} \right] \vee \left[\begin{array}{c} p \equiv \sim q \\ r \end{array} \right] \right] \\
 & \left[\left[\begin{array}{c} p \\ q \end{array} \right] \vee \left[\begin{array}{c} \sim p \\ \sim q \end{array} \right] \right] \vee \left[\left[\begin{array}{c} p \\ \sim q \end{array} \right] \vee \left[\begin{array}{c} \sim p \\ q \end{array} \right] \right]
 \end{aligned}$$

$$[[p \wedge q \wedge \sim r] \vee [\sim p \wedge \sim q \wedge \sim r] \vee [p \wedge \sim q \wedge r] \vee [\sim p \wedge q \wedge r]].$$

Note that we now have three methods for putting a wff into disjunctive normal form:

- (1) Use the method in the proof of 1403 and 1404.

- (2) Put it into nnf and use distributive laws.
- (3) Put it into nnf and use 1407.

Of course, there are close relationships between these methods.

Next we turn our attention to the “method of vertical paths” for seeing that a wff is a contradiction (and hence that it is the negation of a tautology) and the “method of horizontal paths” for seeing that a wff is a tautology. In each case, all one has to do is to look at a vpform for the wff.

DEFINITION. Two literals are *complementary* or *mated* iff one is the negation of the other.

1408 Theorem. Let \mathbf{A} be a wff in negation normal form. Then

- (a) \mathbf{A} is a contradiction iff every path in $\mathcal{VP}(\mathbf{A})$ contains complementary literals.
- (b) \mathbf{A} is a tautology iff every path in $\mathcal{HP}(\mathbf{A})$ contains complementary literals.

Proof of (a): If every path in $\mathcal{VP}(\mathbf{A})$ contains complementary literals, then no assignment can satisfy every literal on any path in $\mathcal{VP}(\mathbf{A})$, so by 1406(a) \mathbf{A} is not satisfiable, i.e., it is a contradiction. If some path in $\mathcal{VP}(\mathbf{A})$ does not contain complementary literals, then by 1406(a) \mathbf{A} is satisfied by the assignment which makes every literal in that vertical path true, so \mathbf{A} is not a contradiction.

Proof of (b): If every path in $\mathcal{HP}(\mathbf{A})$ contains complementary literals, then no assignment can falsify every literal on any path in $\mathcal{HP}(\mathbf{A})$, so by 1406(b) \mathbf{A} is not falsified by any assignment, so it is a tautology. If some path in $\mathcal{HP}(\mathbf{A})$ does not contain complementary literals, then by 1406(b) \mathbf{A} is falsified by the assignment which makes every literal in that horizontal path false, so \mathbf{A} is not a tautology. ■

EXAMPLE: One can easily see from Figure 1.18 that the vpform for W1401 has vertical paths which do not contain complementary literals, and it has horizontal paths which do not contain complementary literals. Therefore, it is neither a contradiction nor a tautology.

EXAMPLE: Let W1402 be the wff

$$[p \vee q] \wedge \sim [p \vee r] \supset \sim r \wedge \neg s \supset q$$

$$\left[\begin{array}{c} p \vee q \\ \sim p \\ \sim r \\ r \vee \left[\begin{array}{c} s \\ \sim q \end{array} \right] \end{array} \right]$$

Figure 1.19: Vpform for $\sim W1402$.

We use the method of vertical paths to show that $\models W1402$.

The following are equivalent:

$$\begin{aligned} & \sim W1402 \\ & \sim \neg [p \vee q] \wedge \sim [p \vee r] \supset \sim r \wedge \neg s \supset q \\ & [p \vee q] \wedge [\sim p \wedge \sim r] \wedge \neg r \vee s \wedge \sim q \end{aligned}$$

The vpform for this wff is displayed in Figure 1.19. One can see that all its vertical paths contain complementary literals, so by 1408, $\sim W1402$ is a contradiction, so $\models W1402$.

EXAMPLE: Let $W1403$ be the wff $[(p \equiv q) \equiv r] \supset \neg p \equiv (q \equiv r)$. We use the method of vertical paths to show that $\models W1403$.

The following are equivalent:

$$\begin{aligned} & \sim W1403 \\ & [(p \equiv q) \equiv r] \wedge \sim [p \equiv (q \equiv r)] \\ & [(p \equiv q) \equiv r] \wedge [\sim p \equiv (q \equiv r)] \\ & [((p \equiv q) \wedge r) \vee ((p \not\equiv q) \wedge \sim r)] \wedge \neg (\sim p \wedge \neg q \equiv r) \vee (p \wedge \neg q \not\equiv r) \\ & [([(p \wedge q) \vee (\sim p \wedge \sim q)] \wedge r) \vee ([(p \wedge \sim q) \vee (\sim p \wedge q)] \wedge \sim r)] \wedge \\ & \neg (\sim p \wedge \neg (q \wedge r)) \vee (\sim q \wedge \sim r) \vee (p \wedge \neg (q \wedge \sim r)) \vee (\sim q \wedge r) \end{aligned}$$

The vpform for this wff is displayed in Figure 1.20. One can see that all its vertical paths contain complementary literals, so by 1408 $\sim W1403$ is a contradiction, so $\models W1403$.

In summary, to determine whether a wff B is a tautology by the method of vertical paths, one may put $\sim B$ into negation normal form, display it in a

$$\left[\left[\begin{array}{c} p \\ q \end{array} \right] \vee \left[\begin{array}{c} \sim p \\ \sim q \end{array} \right] \right] \vee \left[\begin{array}{c} p \\ \sim q \end{array} \right] \vee \left[\begin{array}{c} \sim p \\ q \end{array} \right]$$

$$\left[\left[\begin{array}{c} \sim p \\ r \end{array} \right] \vee \left[\begin{array}{c} \sim q \\ \sim r \end{array} \right] \right] \vee \left[\begin{array}{c} \sim r \\ p \end{array} \right] \vee \left[\begin{array}{c} q \\ \sim r \end{array} \right] \vee \left[\begin{array}{c} \sim q \\ r \end{array} \right]$$

Figure 1.20: Vpform for $\sim W1403$.

vpform, and check whether each vertical path through this vpform contains complementary literals.

We conclude this section with a list of basic tautologies which the reader should verify and become familiar with.

BASIC TAUTOLOGIES

LAWS OF NEGATION

$\sim\sim p$	\equiv	p	(Law of double negation)
$\sim [p \wedge q]$	\equiv	$\neg\neg p \vee \neg q$	(De Morgan's Law)
$\sim [p \vee q]$	\equiv	$\neg\neg p \wedge \neg q$	(De Morgan's Law)
$\sim [p \supset q]$	\equiv	$\neg p \wedge \neg q$	
$\sim [p \equiv q]$	\equiv	$\neg\neg p \equiv q$	
$\sim [p \equiv q]$	\equiv	$\neg p \equiv \neg q$	
$\sim p$	\equiv	$\neg p \mid p$	
$\sim p$	\equiv	$\neg p \downarrow p$	

LAWS DEFINING CONNECTIVES

$p \vee q$	\equiv	$\neg\neg p \supset q$
$p \vee q$	\equiv	$\neg p \supset q \supset q$
$p \wedge q$	\equiv	$\neg p \equiv \neg p \supset q$
$p \supset q$	\equiv	$\neg\neg p \vee q$
$p \supset q$	\equiv	$\neg\neg p \wedge \neg q$
$p \supset q$	\equiv	$\neg p \equiv \neg p \wedge q$
$p \equiv q$	\equiv	$\neg[p \supset q] \wedge [q \supset p]$
$p \equiv q$	\equiv	$\neg[\sim p \vee q] \wedge [p \vee \sim q]$
$p \equiv q$	\equiv	$\neg[p \wedge q] \vee [\sim p \wedge \sim q]$
$p \mid q$	\equiv	$\neg\neg p \wedge q$
$p \downarrow q$	\equiv	$\neg\neg p \vee q$

DISTRIBUTIVE LAWS

$$\begin{aligned} p \wedge [q \vee r] &\equiv \neg[p \wedge q] \vee [p \wedge r] \\ p \vee [q \wedge r] &\equiv \neg[p \vee q] \wedge [p \vee r] \end{aligned}$$

COMMUTATIVE LAWS

$$\begin{aligned} p \wedge q &\equiv \neg q \wedge p \\ p \vee q &\equiv \neg q \vee p \\ p \equiv q &\equiv \neg q \equiv p \\ p \neq q &\equiv \neg q \neq p \\ p \mid q &\equiv \neg q \mid p \\ p \downarrow q &\equiv \neg q \downarrow p \end{aligned}$$

ASSOCIATIVE LAWS

$$\begin{aligned} [p \wedge q] \wedge r &\equiv \neg p \wedge \neg q \wedge r \\ [p \vee q] \vee r &\equiv \neg p \vee \neg q \vee r \\ [p \equiv q] \equiv r &\equiv \neg p \equiv \neg q \equiv r \\ [p \neq q] \neq r &\equiv \neg p \neq \neg q \neq r \end{aligned}$$

TRANSITIVE LAWS

$$\begin{aligned} [p \supset q] \wedge [q \supset r] &\supset \neg p \supset r \\ [p \equiv q] \wedge [q \equiv r] &\supset \neg p \equiv r \end{aligned}$$

LAWS OF t AND f

$$\begin{array}{ll} \sim t \equiv f & p \supset t \equiv t \\ \sim f \equiv t & p \supset f \equiv \sim p \\ p \wedge t \equiv p & t \supset p \equiv p \\ p \wedge f \equiv f & f \supset p \equiv t \\ p \vee t \equiv t & p \equiv t \equiv p \\ p \vee f \equiv p & p \equiv f \equiv \sim p \end{array}$$

OTHER LAWS

$\sim p$	\vee	p	(Law of excluded middle)
p	\supset	p	
p	\equiv	p	
p	\equiv	$\neg p \wedge p$	(Idempotent Law)
p	\equiv	$\neg p \vee p$	(Idempotent Law)
p	\equiv	$\neg p \wedge [q \vee p]$	(Absorption Law)
p	\equiv	$\neg p \vee [q \wedge p]$	(Absorption Law)
$p \wedge q$	\equiv	$\neg p \wedge [q \vee \neg p]$	
$p \vee q$	\equiv	$\neg p \vee [q \wedge \neg p]$	
$p \equiv q$	\equiv	$\neg \sim p \equiv \sim q$	
$p \supset q$	\equiv	$\neg \sim q \supset \sim p$	(Contrapositive law)
$p \wedge q \supset r$	\equiv	$\neg p \supset \neg q \supset r$	
$[p \supset q] \wedge [\sim p \supset r]$	\equiv	$\neg [p \wedge q] \vee [\sim p \wedge r]$	
$p \supset q \supset p$	\supset	p	(Peirce's law)
p	\supset	$\neg p \vee q$	
$p \wedge q$	\supset	p	

EXERCISES

X1400. Prove that the only binary connectives which, taken alone, constitute complete sets of connectives are $|$ and \downarrow .

X1401. Prove that $\{\sim, \equiv\}$ is not a complete set of connectives . (*Hint:* can you find a wff constructed from these connectives and containing exactly two propositional variables which is true in an odd number of lines of the truth table?)

X1402. Let \mathcal{P}' be a system obtained from \mathcal{P} by adding propositional constants t and f to the list of proper symbols of \mathcal{P} . The formation rules of \mathcal{P}' are those of \mathcal{P} plus the additional rule: (0) t and f are wffs. The axiom schemata \mathcal{P}' are the same as those of \mathcal{P} , but of course there are certain instances of these schemata which are axioms of \mathcal{P}' but not wffs (and hence not axioms) of \mathcal{P} . In addition, \mathcal{P}' has the wffs t and $\sim f$ as axioms. For any assignment φ , $\mathcal{V}_\varphi t = T$ and $\mathcal{V}_\varphi f = F$. Show that every tautology of \mathcal{P}' is a theorem of \mathcal{P}' .

X1403. Show that if A is a wff in which no propositional connective other than \equiv occurs, and φ is an assignment, then $\mathcal{V}_\varphi A = F$ iff the number of occurrences of variables p in A such that $\varphi p = F$ is odd. (*Hint:* Consider some examples.)

X1404. Prove that a wff of propositional calculus containing only the connective \equiv is a tautology iff each propositional variable occurs an even number of times. (*Hint:* What other exercise is related to this one?)

X1405. In a wff containing only \wedge , \neq , t , and f as connectives, one may write \wedge as \times , \neq as $+$, t as 1, and f as 0 (since operations on truth values may be regarded as operations of arithmetic modulo 2). Note that in this notation the following wffs are tautologies:

- | | |
|---|-----------------------------|
| (a) $[[P + Q] + R] \equiv \cdot P + [Q + R]$ | (f) $[P + P] \equiv 0$ |
| (b) $[[P \times Q] \times R] \equiv \cdot P \times [Q \times R]$ | (g) $[P \times P] \equiv P$ |
| (c) $[P + Q] \equiv \cdot Q + P$ | (h) $[P + 0] \equiv P$ |
| (d) $[P \times Q] \equiv \cdot Q \times P$ | (i) $[P \times 0] \equiv 0$ |
| (e) $[[P + Q] \times R] \equiv \cdot [P \times R] + [Q \times R]$ | (j) $[P \times 1] \equiv P$ |

Corresponding to each schema of the form $[D \equiv E]$ one has the rule of inference: to replace any occurrence of D in a wff by an occurrence of E . By using rules (a)-(e) one can transform any wff into a polynomial in its propositional variables, which can be further simplified with the aid of rules (f)-(j).

Show that a wff A (in which \times , $+$, 1, and 0 are the only propositional connectives) is a tautology iff there is a sequence of applications of rules (a)-(j) which transforms A into 1.

X1406. Using the notation of X1405, write wffs equivalent to the following wffs, and simplify them:

- | | |
|--|--|
| (a) $p \supset q$ | (d) $[p \equiv q] \wedge [p \vee r] \supset \cdot q \vee r$ |
| (b) $[p \supset q] \wedge [q \supset p]$ | (e) $[p \equiv \cdot q \wedge r] \wedge [q \neq r] \supset \sim p$ |
| (c) $p \vee q$ | (f) $p \wedge [p \supset q] \supset q$ |

Find conjunctive and disjunctive normal forms for the wffs in Exercises X1407–X1409.

X1407. $p \equiv \sim [q \wedge r]$

X1408. $[q \vee r] \supset p$

X1409. $[p \vee r] \equiv [q \wedge \sim p]$

X1410. A binary connective $*$ is *associative* iff $[[p * q] * r] \equiv \cdot p * \cdot q * r$ is a tautology. Is Sheffer's stroke associative?

X1411. Find a wff equivalent to $[p \supset q]$ in which the only connective is Sheffer's stroke.

X1412. Let \mathcal{P}_L be the system of propositional calculus which has Sheffer's stroke as sole connective, and the following axiom and rules of inference:

Axiom. $p \mid [q \mid r] \mid \neg p \mid [r \mid p] \mid \neg s \mid q \mid \neg p \mid s \mid \neg p \mid s$

Rules of Inference.

(1) (Substitution) From \mathbf{A} infer $S_B^P \mathbf{A}$.

(2) From $\mathbf{A} \mid \neg \mathbf{B} \mid \mathbf{C}$ and \mathbf{A} infer \mathbf{C} .

Show that \mathcal{P}_L is consistent, in the sense that there is no wff \mathbf{A} such that both \mathbf{A} and $\mathbf{A} \mid \mathbf{A}$ are theorems.

X1413. Let \mathcal{P}_2 be a formulation of propositional calculus in which the primitive connectives are \wedge , \vee , and \sim . Without using 1404, give a direct inductive proof that if \mathbf{A} is any wff of \mathcal{P}_2 , there is a wff \mathbf{A}' of \mathcal{P}_2 in negation normal form such that $\models \mathbf{A} \equiv \mathbf{A}'$.

X1414. Is $\{\equiv, \supset, \wedge, \vee\}$ a complete set of propositional connectives? Justify your answer.

X1415. For each wff below, find an equivalent wff which contains no connectives other than \sim , \vee , \wedge , \supset , and \equiv , and which contains as few occurrences of propositional variables as possible.

- | | |
|--|---|
| (a) $p \supset q \supset q$ | (d) $[p \wedge q] \vee [p \wedge r]$ |
| (b) $[p \vee q] \wedge p$ | (e) $p \equiv \sim[q \equiv p]$ |
| (c) $\sim[\sim p \wedge \neg p \supset q]$ | (f) $\sim[q \supset \sim \neg q \supset q]$ |

In exercises X1416–X1417 a wff W is given. In each case, do the following:

- Find a wff in negation normal form which is equivalent to $\sim W$.
- Display your answer to part (a) in a vpform.
- Decide whether W is a tautology. If it is not, give an assignment which falsifies it.

X1416. $[[p \wedge r] \vee [q \wedge \sim r]] \supset \neg[q \supset \neg r \wedge s] \supset r$

X1417. $[[p \wedge q] \vee [\sim p \wedge r]] \equiv \neg[p \supset q] \wedge [\sim p \supset r]$

X1418. Prove that $[\lambda p_1 \dots \lambda p_n \mathbf{A}]$ and $[\lambda p_1 \dots \lambda p_n \mathbf{B}]$ denote the same truth function iff $[\mathbf{A} \equiv \mathbf{B}]$ is a tautology.

§15. Compactness

In §12 we explored one aspect of the relationship between the semantics and syntax of \mathcal{P} by showing that the tautologies are the same as the theorems. In this section we show that satisfiability and consistency are coextensive.

1500 Proposition. Let \mathcal{H} be a finite set of wffs of \mathcal{P} . \mathcal{H} is satisfiable iff \mathcal{H} is consistent in \mathcal{P} .

Proof: Suppose \mathcal{H} is not satisfiable. Since the empty set of wffs is trivially satisfiable, \mathcal{H} is nonempty, so we can write \mathcal{H} as $\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$. Since \mathcal{H} is not satisfiable, $\mathbf{A}_1 \wedge \dots \wedge \mathbf{A}_n$ is a contradiction, so for any wff \mathbf{B} , $\models \mathbf{A}_1 \supset \neg \mathbf{A}_2 \supset \dots \supset \neg \mathbf{A}_n \supset \mathbf{B}$. By 1204 this tautology is a theorem, so $\mathcal{H} \vdash \mathbf{B}$ by Modus Ponens. Since \mathbf{B} is an arbitrary wff, \mathcal{H} is inconsistent.

Suppose \mathcal{H} is satisfiable, but inconsistent. Then $\mathcal{H} \vdash p \wedge \neg p$. Let φ be an assignment which satisfies \mathcal{H} . Modus Ponens preserves satisfaction by φ (see the proof of 1200), so $\models_{\varphi} p \wedge \neg p$. However, $p \wedge \neg p$ is not satisfiable. Therefore \mathcal{H} must be consistent. ■

We shall show that 1500 is also true when \mathcal{H} is infinite. The main part of this argument is supplied by the Compactness Theorem below, which is a purely semantic theorem, and does not depend on the details of any axiomatic treatment.

1501 Compactness Theorem. Let \mathcal{S} be an infinite set of wffs of propositional calculus such that every finite subset of \mathcal{S} is satisfiable. Then \mathcal{S} is satisfiable.

Proof: Let $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots$ be the variables which occur in wffs of \mathcal{S} . We shall define an assignment θ by defining $\theta \mathbf{p}_i$ by induction on i . Simultaneously we shall establish by complete induction on i that for every finite subset \mathcal{H} of \mathcal{S} , θ is extendable to \mathcal{H} , i.e., there is an assignment φ which satisfies \mathcal{H} such that $\varphi \mathbf{p}_j = \theta \mathbf{p}_j$ for all $j \leq i$.

Suppose this has been done for all $k < i$. If for every finite subset \mathcal{H} of \mathcal{S} there is an assignment φ which satisfies \mathcal{H} such that $\varphi \mathbf{p}_j = \theta \mathbf{p}_j$ for all $j < i$ and $\varphi \mathbf{p}_i = \top$, let $\theta \mathbf{p}_i = \top$ and we are done at stage i . Otherwise, let $\theta \mathbf{p}_i = \perp$.

In the latter case there is a finite subset \mathcal{K} of \mathcal{S} such that for every assignment φ satisfying \mathcal{K} , if $\varphi \mathbf{p}_j = \theta \mathbf{p}_j$ for all $j < i$, then $\varphi \mathbf{p}_i = \perp$. Let \mathcal{H} be any finite subset of \mathcal{S} . Then $\mathcal{K} \cup \mathcal{H}$ is also a finite subset of \mathcal{S} , so by inductive hypothesis (or by the hypothesis of the theorem when $i = 1$) there

is an assignment φ satisfying $\mathcal{K} \cup \mathcal{H}$ such that $\varphi p_j = \theta p_j$ for all $j < i$. φ satisfies \mathcal{K} as well as \mathcal{H} , so $\varphi p_i = F = \theta p_i$. Thus the required assignment φ exists.

Finally we prove that θ satisfies \mathcal{S} . Let A be any wff in \mathcal{S} , and let i be the largest index of a variable p_i which occurs in A . There is an assignment φ which satisfies $\{A\}$ such that $\varphi p_j = \theta p_j$ for all $j \leq i$. Thus φ and θ agree on all variables which occur in A , so $\mathcal{V}_\theta A = \mathcal{V}_\varphi A = T$. ■

REMARK. Sometimes one gains insight into an argument by considering a situation to which it does *not* apply. Suppose in some generalized system there are “wffs” A such that $\mathcal{V}_\varphi A$ depends on the values that φ assigns to infinitely many variables. For example, suppose $\mathcal{S} = \{A_i \mid i = 1, 2, 3, \dots\}$ where $\mathcal{V}_\varphi A_i = T$ iff $\varphi p_i = T$ and there is a $j > i$ such that $\varphi p_j = F$. (Informally we may write $A_i = [p_i \wedge (\exists j > i) \sim p_j]$. Of course, this is not a wff of propositional calculus.)

It is easy to see that every finite subset of \mathcal{S} is satisfiable, but \mathcal{S} is not satisfiable. Thus it appears that the Compactness Theorem depends in an essential way on the fact that each wff of propositional calculus contains only finitely many propositional variables.

1502 Corollary Let \mathcal{S} be any set of wffs of \mathcal{P} . \mathcal{S} is satisfiable iff \mathcal{S} is consistent in \mathcal{P} .

Proof: (See Figure 1.21.) If \mathcal{S} is inconsistent, then there is a finite subset \mathcal{H} of \mathcal{S} such that $\mathcal{H} \vdash p \wedge \sim p$, so \mathcal{H} is unsatisfiable by 1500, so \mathcal{S} is unsatisfiable too.

Suppose \mathcal{S} is consistent. Then every finite subset \mathcal{H} of \mathcal{S} is consistent, and hence satisfiable by 1500. Therefore \mathcal{S} is satisfiable by 1501. ■

$$\begin{array}{ccc}
 \mathcal{S} \text{ is consistent} & \xrightleftharpoons[1502]{\iff} & \mathcal{S} \text{ is satisfiable} \\
 \text{trivial } \Updownarrow & & 1501 \uparrow\Downarrow \text{ trivial} \\
 \text{Every finite subset} & \xrightleftharpoons[1500]{\iff} & \text{Every finite subset} \\
 \text{of } \mathcal{S} \text{ is consistent} & & \text{of } \mathcal{S} \text{ is satisfiable}
 \end{array}$$

Figure 1.21: Proof of 1502

APPLICATION: We shall show how the Compactness Theorem can be used to establish that every partial ordering of an infinite (but countable) set can be extended to a total ordering. We shall use the fact that every partial ordering of a finite set can be extended to a total ordering; this can easily be proved by induction on the size of the set.

A *partial ordering* of a set S is a relation on S which is reflexive, transitive, and antisymmetric. A set which is endowed with a specified partial ordering is often called a *poset* (partially ordered set). A partial ordering \preceq on a set S is a *total ordering* or *linear ordering* iff for all x and y in S , $x \preceq y$ or $y \preceq x$. To say that a partial ordering \preceq_1 on S can be extended to a linear ordering means that there is a linear ordering \preceq_2 on S such that $x \preceq_2 y$ whenever $x \preceq_1 y$.

Now let a partial ordering \preceq_1 on a countable infinite set S be given. For each ordered pair $\langle x, y \rangle$ of elements from S , specify a propositional variable p_{xy} so that distinct pairs correspond to distinct propositional variables. (We shall think of p_{xy} as meaning that $x \preceq y$.) Let \mathcal{W} be the set of all wffs with one of the following forms, where x , y , and z are in S :

- p_{xy} , where $x \preceq_1 y$;
- p_{xx} ¹⁰;
- $[p_{xy} \wedge p_{yz} \supset p_{xz}]$;
- $\sim [p_{xy} \wedge p_{yx}]$, where $x \neq y$;
- $[p_{xy} \vee p_{yx}]$.

To apply the Compactness Theorem, let any finite subset \mathcal{W}' of \mathcal{W} be given. Let S' be the finite subset of S consisting of those elements x such that a variable of the form p_{xy} or p_{yx} (or p_{xx}) occurs in a wff of \mathcal{W}' . The restriction \preceq'_1 of \preceq_1 to S' is a partial ordering of S' , which can be extended to a total ordering \preceq'_2 of S' since S' is finite. Define the assignment φ as follows: $\varphi p_{xy} = T$ if x and y are elements of S' such that $x \preceq'_2 y$; if q is any propositional variable which is not of this form, $\varphi q = F$. It is easy to see that φ satisfies \mathcal{W}' . Thus every finite subset of \mathcal{W} is satisfiable, so by the Compactness Theorem there is an assignment ψ which satisfies \mathcal{W} .

Now define the relation \preceq_2 on S as follows: $x \preceq_2 y$ iff $\psi p_{xy} = T$. Since ψ satisfies \mathcal{W} , it is easy to see that \preceq_2 is an extension of \preceq_1 which is reflexive, transitive, antisymmetric, and is a total ordering of S .

¹⁰There is admittedly redundancy here, since wffs of this form also have the form above, but it is important to note that wffs of this form are in \mathcal{W} .

EXERCISES

X1500. Extend the proof of the Compactness Theorem so that it will apply to a formulation of the propositional calculus in which there are uncountably many propositional variables. Assume that the propositional variables are well ordered.

X1501. A set \mathcal{S} of wffs of propositional calculus is *disjunctively valid* iff for every assignment φ , there exists $A \in \mathcal{S}$ such that $\models_{\varphi} A$. Prove that a set of wffs is disjunctively valid iff some finite subset of it is disjunctively valid.

X1502. Let \mathcal{T} be the set of all assignments of truth values to propositional variables of \mathcal{P} . For any nonempty set \mathcal{S} of wffs, let $\text{Sat}(\mathcal{S}) = \{\varphi \in \mathcal{T} \mid (\forall A \in \mathcal{S}) \models_{\varphi} A\}$, i.e., $\text{Sat}(\mathcal{S})$ is the set of assignments which satisfy every wff in \mathcal{S} . A subset \mathcal{C} of \mathcal{T} is said to be *closed* iff $\mathcal{C} = \text{Sat}(\mathcal{S})$ for some nonempty set \mathcal{S} of wffs.

- (1) Prove that every finite union of closed sets is a closed set.
- (2) Prove that an arbitrary intersection of closed sets is a closed set.
- (3) Prove that \mathcal{T} is closed.

Thus \mathcal{T} may be regarded as a topological space . A family \mathcal{F} of subsets of \mathcal{T} is said to have the *finite intersection property* iff the intersection of each finite subfamily of \mathcal{F} is nonempty, and \mathcal{T} is said to be *compact* iff each family of closed sets which has the finite intersection property has a nonempty intersection.

- (4) Use 1501 to show that \mathcal{T} is compact.

§16. Ground Resolution

Although it is in principle a routine problem to determine whether a wff is a tautology or a contradiction by constructing its truth table, this can be an extremely lengthy task if the wff contains a large number of propositional variables. Therefore, additional approaches are needed. In §14 we discussed the method of vertical paths. In this section we shall discuss the surprisingly useful “ground resolution” method for determining whether a wff in conjunctive normal form is a contradiction. Both of these methods gain added significance from the fact that they can be extended to theorem proving procedures for first- and higher-order logic.

If **C** and **D** are disjunctions of literals, we say that **C** and **D** are *equivalent as clauses* iff one can be transformed into the other by using the substitutivity of equivalence and the associative, commutative, and idempotent laws of disjunction ($([A \vee B] \vee C) \equiv A \vee [B \vee C]$, $A \vee B \equiv B \vee A$, and $A \vee A \equiv A$). Thus $[[p \vee r] \vee \sim q]$ and $[[r \vee [\sim q \vee r]] \vee [p \vee p]]$ and $[r \vee [\sim q \vee [p \vee \sim q]]]$ are all equivalent to each other. Note that **C** and **D** are equivalent as clauses iff they contain the same literals, so if **C** and **D** are not tautologies, they are equivalent as clauses iff $[C \equiv D]$ is a tautology. A *clause* is an equivalence class of disjunctions of zero or more literals. There is a natural one-to-one correspondence between clauses and sets of literals; namely, a clause corresponds to the set of literals which occur in any wff of that clause. Thus one may informally think of a clause as a disjunction of literals or as a set of literals. We shall denote a clause by any of its wffs. Thus we may say that the clause $[r \vee [\sim q \vee [p \vee \sim q]]]$ is the same as the clause $p \vee r \vee \sim q$. We denote the empty clause, which is always false, by \square . The disjunction of two clauses is the equivalence class of disjunctions of wffs from the given clauses. It corresponds to the union of the associated sets of literals.

In some contexts it is natural to think of a clause as a wff (rather than as a class of equivalent wffs), and to apply to clauses terminology which is appropriate to wffs. We shall say that a set of clauses is *contradictory*, *satisfiable*, etc., iff the corresponding set of wffs is. Note that a finite set of clauses is contradictory iff the conjunction of the clauses (regarded as wffs) is a contradiction. Thus a finite set of clauses may be regarded as representing the conjunction of these clauses. Hence every wff in conjunctive normal form is represented by a finite set of clauses.

We may write the disjunction **A** \vee **B** of clauses **A** and **B** as **A** $\dot{\vee}$ **B** when **A** and **B** are disjoint clauses (i.e., clauses with no common literal).

DEFINITION. A clause **C** is a (*ground*) *resolvent* of clauses **A** and **B** iff there is some propositional variable **p** and there are clauses **D** and **E** such that **A** = $p \dot{\vee} D$, **B** = $\sim p \dot{\vee} E$, and **C** = **D** \vee **E**. Note that **C** may be \square if **D** and **E** are both \square .

REMARK. The ground resolvent rule is a special case of a rule known as the *Cut* rule, which is to infer **D** \vee **E** from **M** \vee **D** and $\sim M \vee E$. Of course, when **D** is \square , the Cut rule reduces to Modus Ponens.

DEFINITION. Let \mathcal{C} be a set of clauses, and let **B** be a clause. A *deduction by (ground) resolution of B from \mathcal{C}* is a sequence S_1, \dots, S_n of clauses such that

\mathbf{S}_n is \mathbf{B} and for each i ($1 \leq i \leq n$), \mathbf{S}_i is in \mathcal{C} or \mathbf{S}_i is a resolvent of clauses \mathbf{S}_j and \mathbf{S}_k such that $j < i$ and $k < i$. A *refutation of \mathcal{C} (by resolution)* is a deduction of \square from \mathcal{C} . We shall write $\mathcal{C} \vdash \mathbf{B}$ to mean there is a deduction of \mathbf{B} from \mathcal{C} .

1600 Ground Resolution Theorem. [Robinson, 1965] Let \mathcal{C} be any set of clauses. Then \mathcal{C} is contradictory iff $\mathcal{C} \vdash \square$.

*Proof.*¹¹ Note that $[\mathbf{p} \vee \mathbf{D}] \wedge [\sim p \vee \mathbf{E}] \supset \mathbf{D} \vee \mathbf{E}$ is a tautology even if $\mathbf{D} = \square$ or $\mathbf{E} = \square$. Hence resolution preserves truth (with respect to an assignment), so if φ satisfies \mathcal{C} and $\mathcal{C} \vdash S$ then $\mathcal{V}_\varphi S = \top$. Since $\mathcal{V}_\varphi \square = \mathbb{F}$ for all assignments φ , if $\mathcal{C} \vdash \square$, then \mathcal{C} must be contradictory.

To prove the converse, by the Compactness Theorem (1501) it clearly suffices to suppose that \mathcal{C} is a finite contradictory set of clauses. For any finite nonempty set $\mathcal{C} = \{\mathbf{S}_1, \dots, \mathbf{S}_n\}$ of clauses, define the *number of excess literals in \mathcal{C}* to be $e(\mathcal{C}) = \#(\mathbf{S}_1) + \dots + \#(\mathbf{S}_n) - n$, where $\#(\mathbf{S})$ is the number of (distinct) literals in the clause \mathbf{S} . Thus $e(\mathcal{C})$ measures the extent to which \mathcal{C} differs from a set of one-literal clauses. The proof is by complete induction on $e(\mathcal{C})$.

If $e(\mathcal{C}) \leq 0$, then $\square \in \mathcal{C}$ or every clause in \mathcal{C} is a one-literal clause. In the latter case since \mathcal{C} is contradictory there must be a contradictory pair \mathbf{p} and $\sim \mathbf{p}$ in \mathcal{C} , from which one can derive \square directly.

Next suppose that we are given a contradictory set \mathcal{C} with $e(\mathcal{C}) > 0$, and by inductive hypothesis $\mathcal{D} \vdash \square$ for every contradictory set \mathcal{D} such that $e(\mathcal{D}) < e(\mathcal{C})$.

If \mathcal{C} contains \square we are done. Otherwise \mathcal{C} must contain some clause \mathbf{A} with at least two literals, so we may write $\mathbf{A} = \mathbf{B} \dot{\vee} \mathbf{L}$ with \mathbf{L} a literal and \mathbf{B} a nonempty clause. Let $\mathcal{S} = \mathcal{C} - \{\mathbf{A}\}$ so $\mathcal{C} = \mathcal{S} \cup \{\mathbf{B} \vee \mathbf{L}\}$. Let $\mathcal{C}_1 = \mathcal{S} \cup \{\mathbf{B}\}$ and $\mathcal{C}_2 = \mathcal{S} \cup \{\mathbf{L}\}$. It is easy to see that \mathcal{C}_1 and \mathcal{C}_2 are contradictory, for if φ satisfies \mathcal{C}_2 (for example), then φ satisfies \mathcal{S} and $\mathcal{V}_\varphi \mathbf{L} = \top$ so $\mathcal{V}_\varphi [\mathbf{B} \vee \mathbf{L}] = \top$ so φ satisfies \mathcal{C} . Also $e(\mathcal{C}_1)$ and $e(\mathcal{C}_2)$ are less than $e(\mathcal{C})$ (even if $\mathbf{B} \in \mathcal{S}$ or $\mathbf{L} \in \mathcal{S}$), since $e(\mathcal{C}) = e(\mathcal{S}) + \#(\mathbf{B})$. Therefore by inductive hypothesis $\mathcal{C}_1 \vdash \square$ and $\mathcal{C}_2 \vdash \square$.

We next show that $\mathcal{C} \vdash \square$ or $\mathcal{C} \vdash \mathbf{L}$. Let $\mathbf{A}_1, \dots, \mathbf{A}_m$ be a deduction of \square from \mathcal{C}_1 . We prove by induction on i that $\mathcal{C} \vdash \mathbf{A}_i$ or $\mathcal{C} \vdash \mathbf{A}_i \vee \mathbf{L}$ for $1 \leq i \leq m$.

Case 1. $\mathbf{A}_i \in \mathcal{C}_1$. Then $\mathbf{A}_i \in \mathcal{S} \subseteq \mathcal{C}$, or $\mathbf{A}_i = \mathbf{B}$ so $\mathbf{A}_i \vee \mathbf{L} = \mathbf{B} \vee \mathbf{L} \in \mathcal{C}$.

¹¹This proof is taken from [Anderson and Bledsoe, 1970].

Case 2 \mathbf{A}_i is a resolvent of \mathbf{A}_j and \mathbf{A}_k , where $j < i$ and $k < i$. Thus $\mathbf{A}_j = \mathbf{D}_j \dot{\vee} \mathbf{p}$, $\mathbf{A}_k = \mathbf{D}_k \dot{\vee} \sim \mathbf{p}$, and $\mathbf{A}_i = \mathbf{D}_j \vee \mathbf{D}_k$. \mathbf{L} might be \mathbf{p} , or a literal in \mathbf{D}_j or \mathbf{D}_k , but in all cases the result of resolving \mathbf{A}_j or $\mathbf{A}_j \vee \mathbf{L}$ with \mathbf{A}_k or $\mathbf{A}_k \vee \mathbf{L}$ is seen to be \mathbf{A}_i or $\mathbf{A}_i \vee \mathbf{L}$. Hence using the inductive hypothesis for \mathbf{A}_j and \mathbf{A}_k we see that $\mathcal{C} \vdash \mathbf{A}_i$ or $\mathcal{C} \vdash \mathbf{A}_i \vee \mathbf{L}$.

If $\mathcal{C} \vdash \square$ we are done. Otherwise $\mathcal{C} \vdash \mathbf{L}$ and $\mathcal{S} \subseteq \mathcal{C}$ and $\mathcal{S} \cup \{\mathbf{L}\} \vdash \square$ so $\mathcal{C} \vdash \square$. ■

EXAMPLE: Use the resolution method to show that the following wff is a tautology: $[p \supset \neg q \wedge \neg r \vee s] \wedge [\neg q \vee \neg r] \supset \neg p \supset s$.

Solution. The negation of the given wff \mathbf{A} must be put into conjunctive normal form. We do this in steps as follows:

- (B) $[\neg p \vee \neg q \wedge \neg r \vee s] \wedge [\neg q \vee \neg r] \wedge \neg \neg p \supset s$.
- (C) $[\neg p \vee q] \wedge [\neg p \vee r \vee s] \wedge [\neg q \vee \neg r] \wedge p \wedge \neg s$

The wff \mathbf{C} is equivalent to $\neg \mathbf{A}$, so \mathbf{A} is a tautology iff \mathbf{C} is contradictory, which is the case iff \square can be derived by resolution from the following five clauses:

- (1) $\neg p \vee q$
- (2) $\neg p \vee r \vee s$
- (3) $\neg q \vee \neg r$
- (4) p
- (5) $\neg s$

This we proceed to do as follows:

(6)	q	1,4
(7)	$\neg r$	3,6
(8)	$r \vee s$	2,4
(9)	s	7,8
(10)	\square	5, 9

EXERCISES

X1600. Show that the following wff is contradictory:

$$\begin{aligned}
 & [\sim d \vee \sim e] \wedge [s \vee q] \\
 & \wedge [h \vee \neg k \wedge [m \vee e \vee a] \wedge [p \vee j \vee b]] \\
 & \wedge [[\sim s \wedge \sim q] \vee \neg \sim h \wedge [\sim r \vee \sim a] \wedge [r \vee a]] \\
 & \wedge [\sim r \vee \neg \sim m \wedge \neg \sim k \vee d \vee a] \\
 & \wedge [r \vee m \vee \neg \sim p \wedge g \wedge \sim c] \\
 & \wedge [\sim a \vee \neg [\sim k \vee \sim m] \wedge [\sim b \vee c] \wedge [\sim j \vee \sim g]].
 \end{aligned}$$

X1601. *Definition:* The clause **A** *subsumes* the clause **B** iff every literal which occurs in **A** also occurs in **B**. (Thus when we think of clauses as sets of literals, **A** subsumes **B** iff **A** \subseteq **B**.)

Suppose \mathcal{S} is a set of clauses which has a refutation of length n , and \mathcal{S}' is a set of clauses such that every clause in \mathcal{S} is subsumed by some clause in \mathcal{S}' . Show that \mathcal{S}' has a refutation of length $\leq n$.

In X1602-X1607, decide whether \square can be derived by resolution from the given clauses. If so, derive it; number the clauses and indicate from which clauses each derived clause is obtained. If \square cannot be derived, prove that it cannot be derived.

X1602.	(1) $p \vee r$	(4) $\sim r \vee \sim s$
	(2) $q \vee s$	(5) $q \vee r$
	(3) $\sim p \vee \sim q$	(6) $p \vee s$

X1603.	(1) $p \vee q$	(5) $q \vee \sim r \vee p$
	(2) $\sim p \vee r \vee \sim q \vee p$	(6) $r \vee q \vee \sim p \vee \sim r$
	(3) $p \vee \sim r$	(7) $\sim r \vee q$
	(4) $q \vee \sim p \vee \sim q$	

X1604.	(1) $\sim q \vee r$	(4) $\sim p \vee \sim q \vee \sim r$
	(2) $\sim r \vee p$	(5) $p \vee q \vee r$
	(3) $\sim p \vee q$	

X1605.	(1) $p \vee q \vee r$	(4) $p \vee \sim s$
	(2) $\sim q \vee s$	(5) $\sim p$
	(3) $\sim r \vee s$	(6) $q \vee \sim r$

- X1606.** (1) $\sim p \vee q \vee \sim s$ (5) $\sim p \vee \sim r \vee \sim q \vee s$
 (2) $p \vee r$ (6) $\sim r \vee q \vee \sim s$
 (3) $\sim q \vee s$ (7) $p \vee \sim r \vee q \vee s$
 (4) $r \vee q \vee s$ (8) $\sim q \vee \sim s$

- X1607.** (1) $p_1 \vee q_1 \vee r_1$ (6) $\sim p \vee \sim r \vee \sim q_1$
 (2) $r \vee p_1 \vee s$ (7) $\sim q \vee \sim p \vee \sim q_1 \vee \sim s$
 (3) $\sim p_1$ (8) $\sim p \vee \sim q \vee \sim r \vee \sim r_1$
 (4) p (9) $\sim p \vee \sim r_1 \vee \sim s$
 (5) $\sim p \vee q$

In X1608-X1610, convert the negation of the given wff to a set of clauses and try to derive \square . If you fail to derive \square , show that the given wff is not a tautology.

X1608. $[p \wedge \neg q \supset r] \supset \neg[s \vee p] \wedge \neg \sim r \supset \neg q$

X1609. $[[p \vee q] \wedge [\neg p \vee \neg \sim r \supset s] \wedge [r \supset s] \wedge [r \vee \neg q]] \supset \neg q \wedge s$

X1610. $[[p \vee q] \wedge [\neg s \supset \neg p \supset r] \wedge [[r \vee s] \supset q]] \supset q$

X1611. Suppose we modify the definition of *ground resolvent* to read as follows: a clause **C** is a ground resolvent of clauses **A** and **B** iff there is some propositional variable **p** such that **A** = $\neg p \vee C$ and **B** = p . With this modified definition is Theorem 1600 still provable? Justify your answer.

Chapter 2

First-Order Logic

We next consider first-order logic, which is also called quantification theory and predicate calculus.¹ Initially we discuss first-order logic without equality, and then we discuss first-order logic with equality in §26.

§20. The Language of \mathcal{F}

Having discussed in Chapter 1 how to combine statements using propositional connectives, we next turn to the problem of formally representing the statements themselves in a way that shows more of their structure. We shall focus on those aspects of the structure of statements which are most important for representing reasoning involving these statements.

Statements make assertions about entities (such as people, animals, plants, colors, qualities, numbers, abstractions, ideas, feelings, etc.), which are designated by nouns and pronouns. For convenience we shall refer to these entities as *objects* or *individuals*, slightly extending the usual meanings of these words in English. In our formal language we shall introduce *terms* (which will be defined below) to denote (serve as names for) these individuals.

Statements can be regarded as asserting that individuals have certain properties, or that individuals are related to each other in certain ways. For example, “Jack will go to the picnic” expresses the assertion that the individual Jack has the property of being an individual who will go to the

¹Frege [Frege, 1879] is generally credited with first developing those aspects of first-order logic which go beyond propositional calculus, though others independently developed similar ideas somewhat later, and Frege’s notation never found general favor. See [Church, 1956, pp. 288–294] for more information about the history of first-order logic.

picnic. “112 is prime” expresses the assertion (which happens to be false) that the individual 112 has the property of being a prime number. “Jill votes for Alice” expresses the assertion that the individuals “Jill” and “Alice” are related by the relation “votes for”. “ $5 < 112$ ” expresses the assertion that the individuals 5 and 112 are related by the relation “ $<$ ”. Another way of saying this is to assert that the relation “ $<$ ” *holds* between 5 and 112.

Since it is sometimes complicated to analyze a sentence in a natural language such as English or Chinese, it is convenient to have a quite simple and uniform syntax to represent the assertion that the entity t has the property P , or that the relation R holds between the entities s and t . We shall use Pt to represent the former, and Rst to represent the latter. (Sometimes it will be convenient to introduce commas or parentheses and write these as $P(t)$ and $R(s, t)$, but the commas and parentheses are in principle unnecessary.) Thus, if we let “Prime” be the property of being a prime number, we can express “112 is prime” as “Prime(112)”. Similarly, we can express “Jill votes for Alice” as “VotesFor(Jill, Alice)”. We call names for properties and relations *predicates*, and they play such a fundamental role in first-order logic that it is sometimes called predicate calculus.

In order to be able to deal adequately with reasoning involving individuals, one must be able to represent statements containing words like “all”, “every”, and “some”. To illustrate the relevant ideas, let us suppose that a class of girls holds an election for class president. Each girl is given a ballot containing the names of all the girls in the class, and is allowed to vote for as many members of the class as she pleases. Suppose we want to express in a formal language the assertion that everyone votes for Alice. One way one might try to formalize this is to introduce a term “everyone” into the formal language, and express the assertion as “VotesFor(everyone, Alice)”. In a similar way, to express the assertion that Jill votes for someone, one might introduce a term “someone” into the formal language, and write “VotesFor(Jill, someone)”. If such statements were permitted in the formal language, it would be very natural to also permit “VotesFor(everyone, someone)”. What should this mean? One obvious possibility is “Everyone votes for someone”. However, there is another possibility: “There is someone for whom everyone votes”. Note that these express quite distinct assertions; the latter statement asserts that everyone votes for the same person, while the former statement does not.

Contemplation of this and other examples (particularly those dealing with more complicated statements) leads to the conclusion that a better way of formally representing assertions involving words like “all”, “every”, and “some” is needed. The solution that has been developed is to use *indi-*

$\text{VotesFor}(\text{Jill}, \text{Alice})$	Jill votes for Alice.
$\forall x \text{ VotesFor}(x, \text{Alice})$	Everyone votes for Alice.
$\forall y \text{ VotesFor}(\text{Jill}, y)$	Jill votes for everyone.
$\forall x \text{ VotesFor}(x, x)$	Everyone votes for herself.
$\forall x \forall y \text{ VotesFor}(x, y)$	Everyone votes for everyone.
$\forall x \exists y \text{ VotesFor}(x, y)$	Everyone votes for someone.
$\exists y \forall x \text{ VotesFor}(x, y)$	There is someone for whom everyone votes.
$\exists x \forall y \text{ VotesFor}(x, y)$	There is someone who votes for everyone.
$\forall y \exists x \text{ VotesFor}(x, y)$	Everyone gets someone's vote.
$\exists x \exists y \text{ VotesFor}(x, y)$	Someone votes for someone.
$\exists x \text{ VotesFor}(x, x)$	Someone votes for herself.
$\exists x \text{ VotesFor}(x, \text{Alice})$	Someone votes for Alice.
$\exists y \text{ VotesFor}(\text{Jill}, y)$	Jill votes for someone.

Figure 2.1: Statements about voting in symbolic form and in English.

vidual variables to refer to arbitrary (unspecified) individuals, and express the assertion that “Everyone votes for Alice” as “For every x , x votes for Alice”, and express the assertion that “Jill votes for someone” as “There is a y such that Jill votes for y . Just as we introduced symbols for propositional connectives, we introduce the symbol \forall to abbreviate the phrase “for every” (or “for all”) and the symbol \exists and write “ $\exists y$ ” to abbreviate the phrase “there is a y such that” (where y is any individual variable). Now we can express “Everyone votes for someone” as “ $\forall x \exists y \text{ VotesFor}(x, y)$ ” and “There is someone for whom everyone votes” as “ $\exists y \forall x \text{ VotesFor}(x, y)$ ”. With this symbolic representation, the distinction between these assertions is made very clear. In Figure 2.1 we list a variety of statements about voting in symbolic form and in English. Note how the symbolic representations of these statements make it easy to focus on the essential logical differences between them.

Expressions of the form $\forall x$ and $\exists x$ are called *quantifiers*. They play such a fundamental role in first-order logic that it is sometimes called quantification theory. The reason we actually use the name “first-order logic” will become evident when we discuss higher-order logic at the beginning of section §50.

Just as \vee can be defined in terms of \sim and \wedge , \exists can be defined in terms of \sim and \forall . We can see this from the following example. Suppose our class of girls decides to vote on some issue, but each girl has the option of voting or abstaining (not voting). If we write “ x votes” as “ $\text{Votes}(x)$ ”, then “ $\sim \text{Votes}(x)$ ” means “ x does not vote” or “ x abstains”. Therefore

$\forall x \sim \text{Votes}(x)$ " means "everyone abstains", and " $\sim \forall x \sim \text{Votes}(x)$ " means "not everyone abstains", which is of course equivalent to the assertion that "someone votes". From this and other examples it can be seen that $\exists x P x$ always has the same logical meaning as $\sim \forall x \sim P x$, provided that there is at least one individual.² This provides a contextual definition of \exists in terms of \sim and \forall . Therefore, for the sake of economy, \exists will not be a primitive symbol of the formal system \mathcal{F} of first-order logic which we shall soon introduce.

Sometimes we wish to refer to individuals by more complex expressions than we have used so far, such as "Jill's father", " $x + 3$ ", or "the value of the diamond". We can do this in a uniform way by having in our language notations for functions which map individuals to individuals. If we let $f(x)$ mean "the father of x ", $g(x, y)$ mean " $x + y$ ", and $V(x)$ mean "the value of x ", then we can represent "Jill's father" as $f(\text{Jill})$, " $x + 3$ " as $g(x, 3)$, and "the value of the diamond" as $V(\text{the diamond})$.

The symbols we will use to refer to individuals, functions, and predicates will be *variables* or *constants*. Variables are used to refer to arbitrary elements of the appropriate type, whereas constants denote particular elements. Thus, variables may occur in quantifiers, but constants may not.

Now we are ready to start describing the system \mathcal{F} of first-order logic. The *primitive symbols* of \mathcal{F} are the following:

- (a) Improper symbols: $[] \sim \vee \wedge$
- (b) Individual variables: $u v w x y z u_1 v_1 w_1 \dots u_2 \dots$
- (c) n -ary function variables: $f^n g^n h^n f_1^n g_1^n h_1^n \dots$ for each natural number $n \geq 1$.
- (d) Propositional variables: $p q r s p_1 q_1 r_1 s_1 \dots$
- (e) n -ary predicate variables: $P^n Q^n R^n S^n P_1^n Q_1^n R_1^n S_1^n \dots$ for each natural number $n \geq 1$.

An n -ary symbol is called *unary* or *monadic* if $n = 1$, and it is called *binary* if $n = 2$. There are denumerably many variables of each type. Individual variables may be regarded as 0-ary function variables, and propositional variables may be regarded as 0-ary predicate variables. In addition, there may be finitely or infinitely many individual, function, propositional, or predicate constants. We also admit systems in which there are no variables of certain types; however, individual variables must always be present, and there must be predicate variables or at least one predicate constant. Each choice of constants and variables determines a different formulation of the

²There is little need for a logical system to discuss nothing, so, as is customary, our logical system will be designed with the tacit assumption that there is at least one individual under discussion.

system \mathcal{F} , of course; thus we use \mathcal{F} ambiguously as the name of any one of a variety of quite similar systems.

The *terms* and *wffs* of \mathcal{F} are defined inductively by the following *formation rules*:

- (a) Each individual variable or constant is a term.
- (b) If t_1, \dots, t_n are terms and f^n is an n -ary function variable or constant, then $f^n t_1 \dots t_n$ is a term.
- (c) If t_1, \dots, t_n are terms and P^n is an n -ary predicate variable or constant, then $P^n t_1 \dots t_n$ is a wff. Also, each propositional variable standing alone is a wff. (Wffs of these forms are called *atomic* or *elementary* wffs.)
- (d) If A is a wff, so is $\sim A$.
- (e) If A and B are wffs, so is $[A \vee B]$.
- (f) If A is a wff and x is an individual variable, then $\forall x A$ is a wff.

Sometimes $\forall x A$ is written as $(\forall x)A$ or simply as $(x)A$. We shall often omit the numerical superscripts from function and predicate symbols in terms and wffs, since the context will show what they must be.

We leave it to the reader (exercises X2006 and X2007) to formulate definitions of the sets of terms and wffs in the style of the definition of the set of wffs in §10 and to formulate and prove analogues of Theorem 1000 (the Principle of Induction on the Construction of a Wff) for terms and wffs of \mathcal{F} .

As in the case of propositional calculus, we shall use A, B, C, D, E , etc., as syntactical variables ranging over wffs. Similarly, we use u, v, w, x, y, z , etc., as syntactical variables for individual variables.

While in first-order logic only individual variables may be quantified, in second-order logic the other types of variables may be quantified too. Since variables designate arbitrary elements of the appropriate type, it is normally appropriate to substitute arbitrary expressions of the appropriate type for variables, under suitable conditions. Such substitution for constants is generally not appropriate since it would have the effect of implying that all elements have the special properties of the element denoted by the constant in question.

For the sake of economy, starting with §25 we shall restrict our attention to formulations of \mathcal{F} in which there are no variables other than individual variables.

We adopt from §10 the conventions for regarding $[A \wedge B]$, $[A \supset B]$, and $[A \equiv B]$ as abbreviations for wffs (now wffs of \mathcal{F}). In addition,

$\exists x A$ shall stand for $\sim \forall x \sim A$.

In addition to statements of the sort we have discussed thus far, we need to be able to symbolically represent statements such as “every natural number is a real number” and “some real numbers are natural numbers”. We shall use \mathbb{N} as a notation for the set of natural numbers, and \mathbb{R} as a notation for the set of real numbers. Associated with every property P is the set S of entities which have property P , and associated with every set S is the property P of being in the set S . Whether one starts with P or with S , one sees that an entity x has property P (i.e., Px) if and only if $x \in S$, so it is sometimes convenient to identify a property P with the set S of entities which have that property, and use the same notation for both. For example, an entity is a natural number if and only if it is in the set of natural numbers, and we may write either $\mathbb{N}x$ or $x \in \mathbb{N}$ to express the assertion that x is a natural number.

We shall express the assertion that “every natural number is a real number” by “for every x in \mathbb{N} , x is a real number”, which we write in symbolic form as $(\forall x \in \mathbb{N})Rx$. Similarly, we express the assertion that “some real numbers are natural numbers” by the statement “there is an x in \mathbb{R} such that x is a natural number”, and write this symbolically as $(\exists x \in \mathbb{R})Nx$. More generally, if A is a statement about x , and P is a property, $(\forall x \in P)A$ means “the statement A is true for every element x having the property P ”, and $(\exists x \in P)A$ means “there exists an element x having the property P such that the statement A is true”.

Let us now observe that these statements can be expressed in terms of the notations we have already introduced in the system \mathcal{F} . We first establish that in every situation, $(\forall x \in P)A$ is true if and only if $\forall x [Px \supset A]$ is true.

Let us start by supposing that $(\forall x \in P)A$ is true, and let x be any entity. If Px is false, then $[Px \supset A]$ is true, but if Px is true, then $x \in P$ so our assumption guarantees that A is true too, so $[Px \supset A]$ is true in this case also. Since we can reach this conclusion for every entity x , we see that $\forall x [Px \supset A]$ is true.

Next we suppose that $\forall x [Px \supset A]$ is true. Let x be any entity in P . Then our assumption guarantees that A is true, so we have established that $(\forall x \in P)A$, completing the proof of our assertion.

In a similar way, we can see that in every situation, $(\exists x \in P)A$ is true if and only if $\exists x [Px \wedge A]$ is true. We first suppose that $(\exists x \in P)A$ is true, and choose some entity x in P such that A . Clearly $[Px \wedge A]$ is true, so there is an entity x such that $[Px \wedge A]$ is true, which means that $\exists x [Px \wedge A]$.

is true. Conversely, if we suppose that $\exists x[Px \wedge A]$ is true, we can choose an entity x such that $[Px \wedge A]$ is true, so $x \in P$, so it is indeed true that $(\exists x \in P)A$.

Motivated by these considerations, we introduce the following additional abbreviations for the system \mathcal{F} . Whenever P is a monadic predicate variable or constant and A is a wff,

$(\forall x \in P)A$ shall stand for $\forall x [Px \supset A]$, and

$(\exists x \in P)A$ shall stand for $\exists x [Px \wedge A]$.

$(\forall x \in P)$ and $(\exists x \in P)$ are called *relativized quantifiers*.

The wff A is the *scope* of the quantifier $\forall x$ in $\forall x A$, of $\exists x$ in $\exists x A$, of $(\forall x \in P)$ in $(\forall x \in P)A$, and of $(\exists x \in P)$ in $(\exists x \in P)A$. We shall continue to use the conventions concerning dots and omissions of brackets introduced in §10, supplemented by the convention that in restoring brackets to an abbreviation of a wff, the quantifiers $\forall x$ and $\exists x$, like the connective \sim , are to be given the smallest possible scope.

EXAMPLE: Suppose we interpret the predicate constants P , T , Z , B , and R so that Px means “ x is a person”, Ty means “ y is a ticket”, Zz means “ z is a prize”, Bxy means “ x buys y ”, and Rxz means “ x receives z ”. Then $\exists y[Ty \wedge Bxy]$ means “ x buys a ticket”, $\exists z[Zz \wedge Rxz]$ means “ x receives a prize”, and

$$\forall x [Px \supset \exists y [Ty \wedge Bxy] \supset \exists z [Zz \wedge Rxz]]$$

means “everyone who buys a ticket receives a prize”. Of course, this sentence could also be written as

$$(\forall x \in P) \bullet (\exists y \in T)Bxy \supset (\exists z \in Z)Rxz.$$

EXAMPLE: Interpret P , T , D , and S so that Px means “ x is a person”, Ty means “ y is trouble”, Kuv means “ u knows v ” and Sxy means “ x has seen y ”. Also, let the individual constant i denote the person who is speaking. Then

$$\sim \exists x [Px \wedge \forall y \bullet Ty \wedge Siy \supset Kxy]$$

may be translated into idiomatic English as “Nobody knows the trouble I’ve seen”. Of course, a more literal translation would be “Nobody knows

all the trouble I've seen", but in the first translation the word "all" is tacitly understood.

When one interprets wffs of first-order logic, one has in mind some set which is called the *universe* or the *domain of individuals*, such as the set of real numbers or the set of all objects on earth, and one interprets a wff of the form $\forall x A$ as meaning " A is true for every element x in the domain of individuals". When this domain is finite and one has names for all the individuals, one can dispense with the quantifier $\forall x$ and express the same idea in terms of conjunctions. For example, suppose there are just three individuals in the domain, and the individual constants, a , b , and c are used as names for them.³ Then $[Pa \wedge Pb \wedge Pc]$ expresses the same idea as $\forall x Px$. Thus we see that \forall can be regarded as a generalized conjunction operator. If the domain has infinitely many individuals, with names a_1, a_2, a_3, \dots , then $\forall x Px$ expresses the idea of the infinite conjunction $[Pa_1 \wedge Pa_2 \wedge Pa_3 \wedge \dots]$. Of course, the latter expression is not a wff of first order logic.

In a similar way, \exists may be regarded as a generalized form of disjunction, and $\exists x Px$ can be written instead of $[Pa \vee Pb \vee Pc]$ in the domain $\{a, b, c\}$. The relationship between \forall and \exists mirrors the relationship between \wedge and \vee . Our definition of $\exists x Px$ as $\sim \forall x \sim Px$ mirrors the generalized DeMorgan Law

$$[Pa \vee Pb \vee Pc] \equiv \sim [\sim Pa \wedge \sim Pb \wedge \sim Pc].$$

This relationship also holds for the relativized quantifiers $(\forall x \in S)$ and $(\exists x \in S)$, and all wffs of the forms $(\exists x \in S)A \equiv \sim (\forall x \in S) \sim A$ will be seen (X2139) to be theorems of \mathcal{F} .

Note that in the domain $\{a, b, c\}$, $\forall x \exists y Rxy$ can also be expressed as

$$\forall x [Rxa \vee Rx b \vee Rx c], \text{ as}$$

$$[\exists y Ray \wedge \exists y Rby \wedge \exists y Rcy], \text{ and as}$$

$$[[Raa \vee Rab \vee Rac] \wedge [Rba \vee Rbb \vee Rbc] \wedge [Rca \vee Rcb \vee Rcc]].$$

The well-formed (wf) parts of a wff B are the consecutive subformulas of B (including B itself) which are wffs.

³The individual constants a , b , and c are symbols of the object language \mathcal{F} , but we shall also use them in our meta-language as names for the individuals in the domain, and thus refer to the domain as $\{a, b, c\}$. Of course, these symbols are also used in the meta-language as names for themselves as symbols of the object language. The context should make clear which usage is intended.

An occurrence of a variable x in a wff B is *bound* iff it is in a wf part of B of the form $\forall x C$; otherwise, it is *free*. For example, in the wff below, the occurrences of variables with “f” above them are free, while those with “b” below them are bound:

$$\begin{array}{ccccccccc} & f & & f & & & & & \\ [\sim [\forall x Pux \vee \forall v \sim \forall u [Pvu \vee \forall u Pux]] \vee \forall y Pyz] & & & & & & & & \\ b & b & b & b & bb & b & b & b & b \end{array}$$

The bound [free] variables of a wff are those which have bound [free] occurrences in the wff. The bound variables in this example are x , v , u , and y , while the free variables are u , x , and z . Note that a variable can be both bound and free in a wff (at different occurrences). A wff or term without free individual variables is said to be *closed*. A closed wff is sometimes called a *cwff*. A *sentence* is a wff without free variables of any type.

Note that a wff can be regarded as a statement about (the objects designated by) its free variables, but that it should not be regarded as asserting anything about its bound variables, which simply serve as pronouns. For example, if the predicate constants C and L are interpreted so that Cy means “ y is a cat” and Lxy means “ x likes y ”, then $\forall y [Cy \supset Lxy]$ means “ x likes all cats”.

If x_1, \dots, x_n are distinct individual variables and t_1, \dots, t_n are terms, the notations $(S_{t_1 \dots t_n}^{x_1 \dots x_n} A)$ and $(S_{\cdot t_1 \dots t_n}^{x_1 \dots x_n} A)$ shall denote, respectively, the results of simultaneously substituting t_i for x_i (for each i such that $1 \leq i \leq n$) at all, and at all free, occurrences of x_i in A . (The parentheses will generally be omitted when the context makes them unnecessary.) The former concept was discussed in §10, and the latter may be defined inductively as follows: Let θ stand for $S_{\cdot t_1 \dots t_n}^{x_1 \dots x_n}$.

- (a) $\theta A = S_{t_1 \dots t_n}^{x_1 \dots x_n} A$ if A is atomic.
- (b) $(\theta \sim B) = \sim (\theta B)$.
- (c) $\theta [B \vee C] = [(\theta B) \vee (\theta C)]$.

- (d) $\theta \forall y B = \begin{aligned} & \forall y \theta B \text{ if } y \notin \{x_1, \dots, x_n\}; \\ & \forall y S_{\cdot t_1 \dots t_{i-1} t_{i+1} \dots t_n}^{x_1 \dots x_{i-1} x_{i+1} \dots x_n} B \text{ if } y = x_i. \end{aligned}$

$S_{t_1 \dots t_n}^{x_1 \dots x_n}$ and $S_{\cdot t_1 \dots t_n}^{x_1 \dots x_n}$ will be called *proper substitutions*.

Sometimes the following substitution notation is convenient. A wff A is written as $A(x_1, \dots, x_n)$, and then $A(t_1, \dots, t_n)$ is understood to stand for $S_{\cdot t_1 \dots t_n}^{x_1 \dots x_n} A$.

In general, it is appropriate to substitute a term [or wff] t for the free occurrences of an individual [or propositional] variable m in a wff A only if

this can be done without the free variables of t becoming bound by quantifiers of \mathbf{A} . For example, most wffs of the form $\forall x \mathbf{A} \supset S_t^x \mathbf{A}$ (where t is a term) are valid (true in all interpretations). In particular, consider the wff $\forall x \exists y [x < y] \supset \exists y [5 < y]$, where $x < y$ is an abbreviation for Pxy and 5 is an individual constant. This is valid, and in the domain of real numbers it can be interpreted as saying, “If for every number x there is a number y which is greater than x , then there is a number greater than 5”. However, if we let the term t be the variable y instead of the constant 5, we obtain the non-valid wff $\forall x \exists y [x < y] \supset \exists y [y < y]$, which can be interpreted as saying, “If for every number x there is a number y greater than x , then there is a number greater than itself”.

Thus, we are led to formulate the following definition of the phrase “ t is *free for m in \mathbf{A}* ”, which expresses in a precise way the assertion that the free variables of t will not become bound by the quantifiers of \mathbf{A} when t is substituted for the free occurrences of m in \mathbf{A} , and that this substitution is therefore legitimate.

DEFINITION. If m is an individual variable and t is a term, or if m is a propositional variable and t is a wff, we say that t is *free for m in the wff \mathbf{A}* iff no free occurrence of m in \mathbf{A} is in a wf part of \mathbf{A} of the form $\forall y \mathbf{B}$, where y is a free variable of t .

EXERCISES

In X2000-X2001, proceed as follows:

- (a) Copy the wff, and underline the free occurrences of variables in it.
- (b) List the free variables of the wff.
- (c) List the bound variables of the wff.
- (d) Decide whether gx is free for u in the wff.
- (e) Decide whether gx is free for z in the wff.
- (f) Decide whether hxy is free for u in the wff.
- (g) Decide whether u is free for x in the wff.

X2000. $\sim [\forall u Puz \wedge \sim \exists x \cdot Qux \vee \forall z \sim Rxz] \supset \exists z Quz$

X2001. $\forall x [Pzu \vee \exists y \cdot Qxy \wedge \sim \forall u Ruz] \supset \forall z Qzu$

In exercises X2002-X2005, interpret the predicate, function, and individual constants as indicated, and translate the given English expression [wff of \mathcal{F}] into a wff of \mathcal{F} [English expression].

X2002. Mx means “ x is a man”; Wx means “ x is a woman”; Bxy means “ x is better than y ”.

- (a) $\forall x [Mx \supset \exists y . Wy \wedge Byx]$
- (b) $\sim \exists z [Wz \wedge \forall u . Mu \supset Bzu]$

X2003. Sxy means “ x and y are sisters”, Lxy means “ x loves y ”, Px means “ x is a person”, Dx means “ x is a dog”, fx means “the wife of x ”, j means “Jack”.

- (a) Somebody loves Jack, but Jack loves only dogs.
- (b) Jack’s wife has a sister who doesn’t love anyone.

X2004. Hx means “ x is a horse”, tx means “the tail of x ”, mx means “the mane of x ”, Bx means “ x is in the barn”, Wy means “ y is white”, Ky means “ y is black”, Lxy means “ x likes y ”.

- (a) Every horse in the barn which has a white tail has a black mane.
- (b) White horses don’t like horses with black manes.
- (c) No horse in the barn has a white tail.
- (d) $(\exists x \in H) [Bx \wedge (\forall y \in H) \neg By \wedge Kty \supset Lxy]$
- (e) $\sim (\exists x \in H) [Bx \wedge \sim Wtx]$

X2005. Bxy means “ x is a brother of y ”, r means “Robert”, Uxy means “ x is an uncle of y ”, f^1x means “the father of x ”, Cxy means “ x is a cousin of y ”, Yxy means “ x is younger than y ”.

- (a) Any brother of Robert’s father is an uncle of Robert.
- (b) $\forall x \forall y [Bf^1x f^1y \supset Cxy]$
- (c) Robert has a cousin who is younger than one of Robert’s brothers.

X2006. Formulate a definition of the set of terms of \mathcal{F} in the style of the definition of the set of wffs of \mathcal{P} in §10, and state and prove an analogue of Theorem 1000.

X2007. Formulate a definition of the set of wffs of \mathcal{F} in the style of the definition of the set of wffs of \mathcal{P} in §10, and state and prove an analogue of Theorem 1000.

X2008. Show that every wf part of a wff $\sim B$ or $\forall x B$ is either the entire wff or a wf part of B .

X2009. Show that every wf part of a wff $[B \vee C]$ is either the entire wff or a wf part of B or a wf part of C . Would this still be true if brackets were omitted from the list of improper symbols and from formation rule (e)?

X2010. Show that if $[B \vee C]$ and $[D \vee E]$ are the same wff, then B is D and C is E .

X2011. Prove by induction on the construction of C that if C is any term or wff and x and y are distinct individual variables such that y is not free in C and y is free for x in C , then $S_x^y S_y^x C = C$. Note explicitly where the conditions on the variables are used in your proof. (*Hint:* when C is $\forall z B$, consider separately the various possibilities concerning z .)

X2012. Let A and C be wffs and let m be a propositional variable. Let x_1, \dots, x_n be the free individual variables of C , and let u_1, \dots, u_n be distinct individual variables which do not occur in C or A . Let B be the wff $S_{u_1 \dots u_n}^{x_1 \dots x_n} C$. Prove that $S_{x_1}^{u_1} \dots S_{x_n}^{u_n} S_B^m A = S_C^m A$. Also give some examples to illustrate this metatheorem, and to show the necessity of the condition on u_1, \dots, u_n .

§21. The Axiomatic Structure of \mathcal{F}

\mathcal{F} has two rules of inference, which are:

Modus Ponens (MP). From A and $A \supset B$ to infer B .

Generalization (Gen). From A to infer $\forall x A$, where x is any individual variable.

We say that we *generalize on x* when we infer $\forall x A$ from A .

The rationale for the Rule of Generalization is that if we are able to establish a statement A containing a free individual variable x , and we have not made any special assumptions about x , then our reasoning about x should apply to every individual in our universe, so the assertion $\forall x A$ should also be true. We shall see that if A does not contain a free occurrence of x , then it is pointless but harmless to generalize on x .

Axiom Schemata for \mathcal{F} .

- (1) $A \vee A \supset A$
- (2) $A \supset \neg B \vee A$
- (3) $A \supset B \supset \neg C \vee A \supset \neg B \vee C$

- (4) $\forall x \mathbf{A} \supset S_t^x \mathbf{A}$ where t is a term which is free for the individual variable x in \mathbf{A} .
- (5) $\forall x [\mathbf{A} \vee \mathbf{B}] \supset \neg \mathbf{A} \vee \forall x \mathbf{B}$ provided that x is not free in \mathbf{A} .

A simple example of Axiom 4 is $[\forall x Px \supset Pa]$, and a simple example of Axiom 5 is $\forall x [p \vee Qx] \supset \neg p \vee \forall x Qx$. In the domain $\{a, b, c\}$ the former wff has the same meaning as $[Pa \wedge Pb \wedge Pc] \supset Pa$, and the latter has the same meaning as

$$[[p \vee Qa] \wedge [p \vee Qb] \wedge [p \vee Qc] \supset \neg p \vee \neg Qa \wedge \neg Qb \wedge \neg Qc].$$

As in the case of the system \mathcal{P} , a *proof* in \mathcal{F} of a wff \mathbf{A} is a sequence of wffs, each of which is an axiom of \mathcal{F} or is inferred from preceding wffs in the sequence by a rule of inference of \mathcal{F} , with \mathbf{A} as the last wff in the sequence. A *theorem* of \mathcal{F} is a wff which has a proof in \mathcal{F} . We now write $\vdash \mathbf{A}$ to mean that \mathbf{A} is a theorem of \mathcal{F} . It is obvious that every wff of \mathcal{P} is a wff of \mathcal{F} , and every theorem of \mathcal{P} is a theorem of \mathcal{F} .

A variety of disciplines can be formalized by adding appropriate additional axioms, often called *postulates*, to those above. These postulates will express the basic definitions and assumptions of the discipline being formalized. Logistic systems obtained in this manner will be called *first-order theories* or *elementary theories*. Normally if the language contains a constant to denote a particular entity of some type, there will also be postulates which express the special properties of that entity.

As an example of a first-order theory we present the elementary theory of additive abelian groups, which we shall call \mathcal{G} . In this theory there are no function or predicate variables, but there is an individual constant 0, a binary function constant +, and a binary predicate constant =. As abbreviations we ordinarily write = ts as $[t = s]$ and +ts as $(t + s)$. The postulates for the theory are the following sentences:

- (G1) $\forall x x = x$
- (G2) $\forall x \forall y \forall z \neg x = y \supset \neg x = z \supset y = z$
- (G3) $\forall x \forall y \forall z \neg x = y \supset x + z = y + z$
- (G4) $\forall x \forall y \forall z \neg x + (y + z) = (x + y) + z$
- (G5) $\forall x \neg x + 0 = x$
- (G6) $\forall x \exists y \neg x + y = 0$
- (G7) $\forall x \forall y \neg x + y = y + x$

DEFINITION. The wff \mathbf{A} is a *substitution instance of a tautology*, or is *tautologous*, iff there is a tautology \mathbf{B} (of \mathcal{P}) such that \mathbf{A} has the form $S_{C_1 \dots C_n}^{P_1 \dots P_n} \mathbf{B}$.

2100 Rule P. If \mathbf{B} is tautologous, then $\vdash \mathbf{B}$. If $\vdash \mathbf{A}_1, \dots, \vdash \mathbf{A}_n$, and if $[\mathbf{A}_1 \wedge \dots \wedge \mathbf{A}_n \supset \mathbf{B}]$ is tautologous, then $\vdash \mathbf{B}$.

Proof: If \mathbf{B} is tautologous, there is a tautology \mathbf{D} such that \mathbf{B} is $S_{C_1 \dots C_n}^{P_1 \dots P_n} \mathbf{D}$. Since \mathbf{D} is a theorem of \mathcal{P} , there is a proof $\mathbf{D}_1, \dots, \mathbf{D}_m$ of \mathbf{D} in \mathcal{P} . It is easily seen by induction on i that $S_{C_1 \dots C_n}^{P_1 \dots P_n} \mathbf{D}_i$ is a theorem of \mathcal{F} for $1 \leq i \leq m$. This proves the first part of the metatheorem.

To prove the second part, note that if $[\mathbf{A}_1 \wedge \dots \wedge \mathbf{A}_n \supset \mathbf{B}]$ is tautologous, then $[\mathbf{A}_1 \supset \bullet \mathbf{A}_2 \supset \bullet \dots \supset \bullet \mathbf{A}_n \supset \mathbf{B}]$ must be tautologous also, so it is a theorem of \mathcal{F} , so by n applications of Modus Ponens $\vdash \mathbf{B}$. ■

REMARK. Rule P is a very powerful derived rule of inference which embodies all the power of reasoning involving propositional connectives. For example, the following are special cases of Rule P:

- (1) **Modus Ponens.** If $\vdash \mathbf{A}$ and $\vdash \mathbf{A} \supset \mathbf{B}$, then $\vdash \mathbf{B}$
(since $[\mathbf{A} \wedge [\mathbf{A} \supset \mathbf{B}] \supset \mathbf{B}]$ is tautologous).
- (2) **Modus Tollens.** If $\vdash \mathbf{A} \supset \mathbf{B}$ and $\vdash \sim \mathbf{B}$, then $\vdash \sim \mathbf{A}$
(since $[\mathbf{A} \supset \mathbf{B}] \wedge \sim \mathbf{B} \supset \sim \mathbf{A}$ is tautologous).
- (3) **Transitive Law of Implication (Trans).** If $\vdash \mathbf{A}_1 \supset \mathbf{A}_2$,
 $\vdash \mathbf{A}_2 \supset \mathbf{A}_3, \dots$, and $\vdash \mathbf{A}_{n-1} \supset \mathbf{A}_n$, then $\vdash \mathbf{A}_1 \supset \mathbf{A}_n$
(since $[\mathbf{A}_1 \supset \mathbf{A}_2] \wedge [\mathbf{A}_2 \supset \mathbf{A}_3] \wedge \dots \wedge [\mathbf{A}_{n-1} \supset \mathbf{A}_n] \supset [\mathbf{A}_1 \supset \mathbf{A}_n]$ is tautologous).

2101 Consistency Theorem. \mathcal{F} is absolutely consistent and consistent with respect to negation.

Proof: For each wff \mathbf{A} of \mathcal{F} , let $\natural \mathbf{A}$ be the result of erasing all quantifiers in \mathbf{A} , and replacing each occurrence of an atomic wff in \mathbf{A} by the propositional variable p . It is easily seen by induction on proofs that if $\vdash \mathbf{A}$, then $\natural \mathbf{A}$ is a tautology. (That is, it is easily seen by induction on i that if $\mathbf{A}_1, \dots, \mathbf{A}_i, \dots, \mathbf{A}_n$ is any proof in \mathcal{F} , then $\natural \mathbf{A}_i$ is a tautology for each i from 1 to n .) However, $\natural \sim \mathbf{A}$ is $\sim \natural \mathbf{A}$, so \mathcal{F} is consistent with respect to negation, and hence, by 1201 and 2100, absolutely consistent. ■

We next prove certain theorem schemata and derived rules of inference of \mathcal{F} . We remark that, in contrast to the situation for \mathcal{P} , the decision problem for \mathcal{F} is unsolvable. (This deep metatheorem is known as *Church's Theorem* [Church, 1936a].) Therefore, techniques for proving theorems of \mathcal{F} , and for showing that certain wffs are not theorems of \mathcal{F} , are of permanent importance. (By contrast, the theorems of \mathcal{P} proved in §11 were simply

needed to prove Theorem 1205, and they are all subsumed by Theorem 1205.) Note that the proof of 2101 provides a technique for showing that certain wffs are not theorems of \mathcal{F} .

2102. $\vdash \forall x_n \dots \forall x_1 A \supset A$.

Proof: By induction on n . For $n = 1$, this is an instance of Axiom Schema 4, since x is always free for x in A . For the induction step, proceed as follows:

- (.1) $\vdash \forall x_{n+1} \forall x_n \dots \forall x_1 A \supset \forall x_n \dots \forall x_1 A$ Axiom 4
- (.2) $\vdash \forall x_n \dots \forall x_1 A \supset A$ by inductive hypothesis
- (.3) $\vdash \forall x_{n+1} \forall x_n \dots \forall x_1 A \supset A$ Rule P: .1, .2

■

2103 $\supset \forall$ Rule. If $\vdash A \supset B$ and x is not free in A , then $\vdash A \supset \forall x B$.

Proof:

- (.1) $\vdash \sim A \vee B$ given
- (.2) $\vdash \forall x \sim A \vee B$ Gen: .1
- (.3) $\vdash \forall x [\sim A \vee B] \supset \sim A \vee \forall x B$ Axiom 5
- (.4) $\vdash A \supset \forall x B$ MP: .2, .3

■

2104. $\vdash \forall x [A \supset B] \supset \forall x A \supset \forall x B$

ILLUSTRATION: If the street gets wet every day that it rains, then if it rains every day, the street gets wet every day.

Proof:

- (.1) $\vdash \forall x [A \supset B] \supset \forall x A \supset B$ 2102
- (.2) $\vdash \forall x A \supset A$ 2102
- (.3) $\vdash \forall x [A \supset B] \wedge \forall x A \supset B$ Rule P: .1, .2
- (.4) $\vdash \forall x [A \supset B] \wedge \forall x A \supset \forall x B$ 2103: .3
- (.5) $\vdash \forall x [A \supset B] \supset \forall x A \supset \forall x B$ Rule P: .4

■

DEFINITION. Let M be a wf part of A . An occurrence of M in A is *positive* [*negative*] in A iff that occurrence is in the scope of an even [odd] number of occurrences of \sim in A . (0 is even).

2105 Substitutivity of Implication. Let M , N , and A be wffs, and let A' be the result of replacing M by N at zero or more occurrences (henceforth called designated occurrences) of M in A . Let y_1, \dots, y_k be a list including all individual variables which occur free in M or N , but are bound in A by quantifiers whose scopes contain designated occurrences of M .

If the designated occurrences of M are all positive in A , then

- (a) $\vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \neg A \supset A'$, and
- (b) if $\vdash M \supset N$, then $\vdash A \supset A'$.

If the designated occurrences of M are all negative in A , then

- (c) $\vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \neg A' \supset A$, and
- (d) if $\vdash M \supset N$, then $\vdash A' \supset A$.

EXAMPLE:

$\vdash \forall y [\forall x Pxy \supset \exists z Qyz] \supset \neg \exists y [\exists z Qyz \supset Ry] \supset \exists y [\forall x Pxy \supset Ry]$ by 2105c.

Proof: Clearly (b) and (d) follow from (a) and (c), respectively, by Gen and MP. We prove (a) and (c) simultaneously by induction on the number of occurrences of \sim , \vee , and \forall in A .

Case 1. There are no designated occurrences of M in A . Then A' is A , so Rule P establishes the result.

Case 2. A is M , and the sole occurrence of M in A is designated. Then M occurs positively in A , and A' is N , so the desired theorem is

$$\forall y_1 \dots \forall y_k [M \supset N] \supset \neg M \supset N,$$

which is an instance of 2102.

In the remaining cases we assume, as part of the definition of these cases, that neither Case 1 nor Case 2 applies.

Case 3. A has the form $\sim B$. Since Cases 1 and 2 do not apply, all designated occurrences in M in A are occurrences in B , and A' is $\sim B'$. In part (a) of the metatheorem the designated occurrences of M are all positive in A , hence negative in B , so

$$(31) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \neg B' \supset B \quad \text{by inductive hypothesis}$$

$$(.32) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet \sim B \supset \sim B' \quad \text{by Rule P}$$

Part (c) is established similarly.

Case 4. A has the form $[B \vee C]$. All designated occurrences of M in A must be in B or in C , and A' is $[B' \vee C']$. Again we establish (a) and leave (c) to the reader. Since the designated occurrences of M are all positive in A , they are positive in B and in C .

$$(.41) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet B \supset B' \quad \text{by inductive hypothesis}$$

$$(.42) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet C \supset C' \quad \text{by inductive hypothesis}$$

(Note that the conditions on the variables $y_1 \dots y_k$ have been stated in such a way that the inductive hypothesis applies to B and C .)

$$(.43) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet [B \vee C] \supset [B' \vee C'] \quad \text{Rule P: .41, .42}$$

Case 5. A is $\forall x B$. All designated occurrences of M in A must be in B , so A' is $\forall x B'$. Note that if x is free in M or in N , then one of the variables y_i must be x , so x is not free in $\forall y_1 \dots \forall y_k [M \supset N]$. Again we treat only part (a).

$$(.51) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet B \supset B' \quad \text{by inductive hypothesis}$$

$$(.52) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \forall x \bullet B \supset B' \quad 2103: .51$$

$$(.53) \vdash \forall x [B \supset B'] \supset \bullet \forall x B \supset \forall x B' \quad 2104$$

$$(.54) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet \forall x B \supset \forall x B' \quad \text{Rule P: .52, .53}$$

Since A must fall under one of these five cases, the proof is complete. ■

2106 Substitutivity of Equivalence. Let M , N , A , A' , and y_1, \dots, y_k be as in 2105.

$$(a) \vdash \forall y_1 \dots \forall y_k [M \equiv N] \supset \bullet A \equiv A'$$

$$(b) \text{ If } \vdash M \equiv N \text{ then } \vdash A \equiv A'.$$

$$(c) \text{ If } \vdash M \equiv N \text{ and } \vdash A, \text{ then } \vdash A'.$$

Proof: Clearly (b) and (c) follow from (a) by Gen and Rule P. To prove (a), let B be the result of replacing all positive designated occurrences of M in A by occurrences of N .

$$(.1) \vdash \forall y_1 \dots \forall y_k [M \supset N] \supset \bullet A \supset B \quad 2105(a)$$

$$(.2) \vdash \forall y_1 \dots \forall y_k [N \supset M] \supset \bullet B \supset A \quad 2105(a)$$

Clearly \mathbf{A}' is the result of replacing certain negative occurrences of \mathbf{M} in \mathbf{B} by occurrences of \mathbf{N} , so

- (.3) $\vdash \forall y_1 \dots \forall y_k [\mathbf{M} \supset \mathbf{N}] \supset \bullet \mathbf{A}' \supset \mathbf{B}$ 2105(c)
- (.4) $\vdash \forall y_1 \dots \forall y_k [\mathbf{N} \supset \mathbf{M}] \supset \bullet \mathbf{B} \supset \mathbf{A}'$ 2105(c)
- (.5) $\vdash [\mathbf{M} \equiv \mathbf{N}] \supset \bullet \mathbf{M} \supset \mathbf{N}$ Rule P
- (.6) $\vdash \forall y_1 \dots \forall y_k [\mathbf{M} \equiv \mathbf{N}] \supset \forall y_1 \dots \forall y_k [\mathbf{M} \supset \mathbf{N}]$ 2105(b): .5
- (.7) $\vdash [\mathbf{M} \equiv \mathbf{N}] \supset \bullet \mathbf{N} \supset \mathbf{M}$ Rule P
- (.8) $\vdash \forall y_1 \dots \forall y_k [\mathbf{M} \equiv \mathbf{N}] \supset \forall y_1 \dots \forall y_k [\mathbf{N} \supset \mathbf{M}]$ 2105(b): .7
- (.9) $\vdash \forall y_1 \dots \forall y_k [\mathbf{M} \equiv \mathbf{N}] \supset \bullet \mathbf{A} \equiv \mathbf{A}'$ Rule P: .1, .2, .3, .4, .6, .8

■

$$\mathbf{2107.} \vdash \forall x [\mathbf{M} \equiv \mathbf{N}] \supset \bullet \forall x \mathbf{M} \equiv \forall x \mathbf{N} \quad 2106(a)$$

■

$$\mathbf{2108.} \vdash \forall x [\mathbf{M} \equiv \mathbf{N}] \supset \bullet \exists x \mathbf{M} \equiv \exists x \mathbf{N} \quad 2106(a)$$

■

2109. $\vdash \forall x \mathbf{C} \equiv \forall y S_y^x \mathbf{C}$, provided that y is not free in \mathbf{C} and y is free for x in \mathbf{C} .

$$\begin{aligned} \text{EXAMPLE: } & \vdash \forall x \exists u [Rxug^1x \vee \forall y \bullet Puy \supset \forall x Px y] \\ & \equiv \forall y \exists u [Ryu g^1y \vee \forall y \bullet Puy \supset \forall x Px y] \end{aligned}$$

ILLUSTRATION: Let us illustrate the restrictions in Theorem 2109 with the following wffs:

- (A) $\forall x [Fx \supset Lzx]$
- (B) $\forall y [Fy \supset Lzy]$
- (C) $\forall z [Fz \supset Lzz]$
- (D) $\forall x [[Px \wedge \exists y \bullet Fy \wedge Hxy] \supset Lzx]$
- (E) $\forall y [[Py \wedge \exists y \bullet Fy \wedge Hyy] \supset Lzy]$

We shall interpret the predicate constants in these wffs so that Fx means “ x is a cat” (or “ x is feline”), Lzx means “ z likes x ”, Px means “ x is a person”, and Hxy means “ x has y ”. Then A and B both mean “ z likes all cats”, but C means “every cat likes itself”. $\vdash A \equiv B$ by 2109, but $A \equiv C$ is not an instance of 2109, since z is free in $[Fx \supset Lzx]$. D means “ z likes everyone who has a cat”, but E means “ z likes every person such that some cat has itself”. $D \equiv E$ is not an instance of 2109, since y is not free for x in $[[Px \wedge \exists y \bullet Fy \wedge Hxy] \supset Lzx]$.

Proof of 2109: We may assume x and y are distinct, since the theorem follows by Rule P if x is y .

$$\begin{array}{ll} (.1) \vdash \forall x C \supset S_y^x C & \text{Axiom 4} \\ (.2) \vdash \forall x C \supset \forall y S_y^x C & \supset \forall \text{ Rule (2103): .1} \end{array}$$

It is easily verified that x is not free in $S_y^x C$ and that x is free for y in $S_y^x C$.

$$(3) \vdash \forall y S_y^x C \supset S_{\dot{x}\dot{y}}^y S_x^x C \quad \text{Axiom 4}$$

Our conditions assure that $S_{\dot{x}\dot{y}}^y S_x^x C$ is C (see Exercise X2011), so

$$(4) \vdash \forall y S_y^x C \supset \forall x C \quad \supset \forall \text{ Rule: .3}$$

The desired result follows by Rule P from .2 and .4. ■

2110 Rule of Alphabetic Change of Bound Variables (α Rule). Suppose that A' is obtained from A upon replacing an occurrence of $\forall x C$ in A by an occurrence of $\forall y S_y^x C$, where y is not free in C and y is free for x in C . If $\vdash A$, then $\vdash A'$.

Proof: By 2109 and the Substitutivity of Equivalence (2106). ■

DEFINITION. A *proof of a wff B from the finite set of hypotheses $\{A_1, \dots, A_n\}$* is a finite sequence B_1, \dots, B_m of wffs such that B_m is B , and for each i one of the following conditions is satisfied:

- (a) (*Ax*) B_i is an axiom;
- (b) (*Hyp*) B_i is a hypothesis A_j ;
- (c) (*MP*) B_i follows by Modus Ponens from preceding members of the sequence;
- (d) (*Gen*) there exists $j < i$ such that B_i is $\forall x B_j$, and x is not free in any hypothesis;
- (e) (α) B_i follows from some preceding member of the sequence by the Rule of Alphabetic Change of Bound Variables.

We write $A_1, \dots, A_n \vdash B$ to indicate that there is a proof of B from the hypotheses A_1, \dots, A_n . If \mathcal{S} is a set of wffs, $\mathcal{S} \vdash B$ means there is a proof of B from some finite subset of \mathcal{S} . It is easily seen that $\vdash B$ iff $\emptyset \vdash B$. As

in Chapter 1 we use \mathcal{H} as a syntactical variable for a set of wffs. When we say x is not free in \mathcal{H} , we shall mean that x is not free in any wff in \mathcal{H} .

The restriction on the Rule of Generalization (Gen) in proofs from hypotheses is a very important and natural one. In general, if we prove a wff B in which x occurs free (which can, of course, be regarded as a statement about x), and have made no special assumptions about x , then the argument should apply to an arbitrary object which could be named by x , so it is intuitively clear that $\forall x B$ should also be true. However, if x occurs free in some hypothesis H , then the proof may depend on the assumption about x which H expresses, so there is no reason to suppose that $\forall x B$ will also be true.

Note that if S is a set of closed wffs which are the postulates for a first-order theory T , then $\vdash_T B$ iff $S \vdash B$ (where $\vdash_T B$ means that B is a theorem of T). Thus, derived rules of inference concerning proofs from hypotheses can also be applied to proofs in first-order theories.

2111 Universal Instantiation ($\forall I$) If $\mathcal{H} \vdash \forall x A$, then $\mathcal{H} \vdash S_t^x A$, provided that t is a term free for x in A . ■

Proof: By Axiom 4 and MP. ■

2112 Lemma. Suppose $\mathcal{H} \vdash A$, and let U be a finite set of individual variables which are not free in A or in \mathcal{H} . Then there is a proof of A from \mathcal{H} in which there is no application of the rule of Generalization involving generalization on a member of U .

Proof: It suffices to prove the lemma for the case where \mathcal{H} is finite. Let R be a proof of A from \mathcal{H} , and let $U = \{x_1, \dots, x_n\}$. Let y_1, \dots, y_n be distinct individual variables which do not occur in U or any wff of R or \mathcal{H} . Let θ be the substitution $S_{y_1 \dots y_n}^{x_1 \dots x_n}$. Clearly θR is a proof of θA from $\theta \mathcal{H}$, and by inserting appropriate alphabetic changes of bound variables it can be expanded to a proof of A from \mathcal{H} . The new proof contains no generalization on any variable in U . ■

2113 Theorem. If $\mathcal{H}_1 \vdash A$ and $\mathcal{H}_1 \subseteq \mathcal{H}_2$, then $\mathcal{H}_2 \vdash A$.

Proof: It suffices to consider the case where \mathcal{H}_1 and \mathcal{H}_2 are both finite. Let y_1, \dots, y_k be the free individual variables of A which are not free in \mathcal{H}_1 and let $\bar{A} = \forall y_1 \dots \forall y_k A$. By Generalization, $\mathcal{H}_1 \vdash \bar{A}$. Let U be the set of individual variables which occur free in \mathcal{H}_2 but not in \mathcal{H}_1 . Then no

member of \mathcal{U} is free in $\overline{\mathbf{A}}$ (since the free variables of $\overline{\mathbf{A}}$ are all free in wffs of \mathcal{H}_1), so by 2112 there is a proof of $\overline{\mathbf{A}}$ from \mathcal{H}_1 in which no member of \mathcal{U} is generalized on. This is a proof from \mathcal{H}_2 , so $\mathcal{H}_2 \vdash \overline{\mathbf{A}}$, and $\mathcal{H}_2 \vdash \mathbf{A}$ by Universal Instantiation. ■

Corollary. If $\vdash \mathbf{A}$, then $\mathcal{H} \vdash \mathbf{A}$.

REMARK. The reader may have wondered why the Rule of Alphabetic Change of Bound Variables occurs in the definition of a proof from hypotheses. It will be observed that this rule was used in the proof of 2112, which was used to prove 2113. Actually, it would not be possible to prove metatheorem 2113 if the α Rule were omitted from the definition of a proof from hypotheses. To see this, note that $\vdash \forall y Py \supset \forall x Px$ (by 2109.2), so $[\mathcal{Q}x \supset \mathcal{Q}x] \vdash \forall y Py \supset \forall x Px$ by 2113. Suppose there were a proof of $\forall y Py \supset \forall x Px$ from $[\mathcal{Q}x \supset \mathcal{Q}x]$ in which the α Rule was not used. Since $[\mathcal{Q}x \supset \mathcal{Q}x]$ can be proved using only Axioms 1–3 and Modus Ponens, it would not actually be necessary to use $[\mathcal{Q}x \supset \mathcal{Q}x]$ as a premise in this proof. Hence the given proof would simply be a proof of $\forall y Py \supset \forall x Px$ in which no generalization is made on the variable x . However, we shall see in §24 that any proof in \mathcal{F} of $[\forall y Py \supset \forall x Px]$ must contain a generalization upon x .

2114 Extended Rule P. If $\mathcal{H} \vdash \mathbf{A}_1, \dots, \mathcal{H} \vdash \mathbf{A}_n$, and if $[\mathbf{A}_1 \wedge \dots \wedge \mathbf{A}_n \supset \mathbf{B}]$ is tautologous, then $\mathcal{H} \vdash \mathbf{B}$. Also, if \mathbf{B} is tautologous, $\mathcal{H} \vdash \mathbf{B}$.

Proof: $\mathcal{H} \vdash \mathbf{A}_1 \supset \mathbf{B} \dots \supset \mathbf{A}_n \supset \mathbf{B}$ by 2100 and 2113. Hence $\mathcal{H} \vdash \mathbf{B}$ by n applications of MP. ■

2115 Extended $\supset \forall$ Rule. If $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}$ and if x is not free in \mathcal{H} or in \mathbf{A} , then $\mathcal{H} \vdash \mathbf{A} \supset \forall x \mathbf{B}$.

The proof is similar to the proof of 2103.

2116 Deduction Theorem (Ded). If $\mathcal{H}, \mathbf{A} \vdash \mathbf{B}$ then $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}$.

Proof. Let $\mathbf{B}_1, \dots, \mathbf{B}_n$ be a proof of \mathbf{B} from \mathcal{H}, \mathbf{A} . We prove that $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}_i$ for $1 \leq i \leq n$ by induction on i .

Case 1. \mathbf{B}_i is an axiom or a member of \mathcal{H} .

(11) $\mathcal{H} \vdash \mathbf{B}_i$	axiom or hypothesis
--	---------------------

(.12) $\mathcal{H} \vdash A \supset B_i$

Rule P: .11

Case 2. B_i is A .

(.21) $\mathcal{H} \vdash A \supset B_i$

Rule P

Case 3. B_i is inferred by MP from B_j and B_k , which is $[B_j \supset B_i]$, where $j, k < i$.

(.31) $\mathcal{H} \vdash A \supset B_j$

by inductive hypothesis

(.32) $\mathcal{H} \vdash A \supset \bullet B_j \supset B_i$

by inductive hypothesis

(.33) $\mathcal{H} \vdash A \supset B_i$

Rule P: .31, .32

Case 4. B_i is inferred from B_j by Gen, where $j < i$. Thus, B_i is $\forall x B_j$, and x does not occur free in \mathcal{H} or in A .

(.41) $\mathcal{H} \vdash A \supset B_j$

by inductive hypothesis

(.42) $\mathcal{H} \vdash A \supset \forall x B_j$

$\supset \forall$ rule (2115): .41

Case 5. B_i is inferred from B_j by the α Rule, where $j < i$.

(.51) $\mathcal{H} \vdash A \supset B_j$

by inductive hypothesis

(.52) $\mathcal{H} \vdash A \supset B_i$

α Rule

■

2117 Extended Substitutivity of Implication and Equivalence. Let M , N , and A be wffs, and let A' be the result of replacing M by N at certain designated occurrences of M in A . If $\mathcal{H} \vdash A$, then $\mathcal{H} \vdash A'$, provided that one of the following conditions is satisfied:

- (a) $\vdash M \supset N$, and the designated occurrences of M are all positive in A ;
- (b) $\vdash N \supset M$, and the designated occurrences of M are all negative in A ;
- (c) $\vdash M \equiv N$.

Proof:

(.1) $\mathcal{H} \vdash A$ given

Case (a).

(.11) $\vdash M \supset N$

given

(.12) $\vdash A \supset A'$

2105(b)

(.13) $\mathcal{H} \vdash A \supset A'$

2113

(14) $\mathcal{H} \vdash A'$

MP: .1, .13

Case (b).(21) $\vdash N \supset M$

given

(22) $\vdash A \supset A'$

2105(d)

(Note that A is obtained from A' upon replacing certain negative occurrences of N in A' by occurrences of M .)(23) $\mathcal{H} \vdash A \supset A'$

2113

(24) $\mathcal{H} \vdash A'$

MP: .1, .23

The proof for Case (c) is like the proof of (a), but uses 2106(b) and Rule P. ■

2118 Rule of Substitution (Sub). Let m be an individual variable and t be a term, or let m be a propositional variable and t be a wff. If $\mathcal{H} \vdash A$, then $\mathcal{H} \vdash S_t^m A$, provided that m does not occur free in \mathcal{H} and t is free for m in A .

Proof: If m is an individual variable, use Gen and $\forall I$ (2111).

If m is a propositional variable, we first consider the case in which \mathcal{H} is empty. Let A_1, \dots, A_k be a proof of A . Let x_1, \dots, x_n be the free individual variables of t , and let u_1, \dots, u_n be distinct individual variables which do not occur in any wff of $\{A_1, \dots, A_k, t\}$. Let B be $S_{u_1 \dots u_n}^{x_1 \dots x_n} t$. It is easy to establish by induction on i that $\vdash S_B^m A_i$ for $1 \leq i \leq k$.

If A_i is an axiom, then $S_B^m A_i$ is too, since no individual variable of A_i occurs free in B . (We leave it to the reader to verify the details when A_i is an instance of Axiom Schema 4 or 5.) The arguments for the cases where A_i is obtained by Modus Ponens or Generalization are routine.

The inductive proof is complete, so $\vdash S_B^m A$. Hence by n applications of the Rule of Substitution for individual variables, $\vdash S_{x_1}^{u_1} \dots S_{x_n}^{u_n} S_B^m A$. (To see that x_i is free for u_i , use the fact that t is free for m in A .) From the definition of B and the conditions on u_1, \dots, u_n it can be seen that $S_{x_1}^{u_1} \dots S_{x_n}^{u_n} B = t$, so since u_1, \dots, u_n do not occur in A , $S_{x_1}^{u_1} \dots S_{x_n}^{u_n} S_B^m A = S_t^m A$ (see Exercise X2012), and the theorem is proved for the case where \mathcal{H} is empty.

Now to prove the theorem for the case that \mathcal{H} is not empty, let $\mathcal{H} = \{H_1, \dots, H_j\}$.

 $H_1, \dots, H_j \vdash A$

given

 $\vdash H_1 \supset \vdash H_2 \supset \dots \vdash H_j \supset A$

Deduction Theorem

$$\vdash H_1 \supset \dots \supset H_j \supset S_t^m A$$

by substitution, since m does not occur in H .

$$\mathcal{H} \vdash S_t^m A$$

2113, MP

■

2119. $\vdash \sim \exists x A \equiv \forall x \sim A$ by Rule P and the Definition of \exists

ILLUSTRATION: Nobody votes if and only if everyone abstains.

2120. $\vdash \sim \forall x A \equiv \exists x \sim A$

by Rule P, Substitutivity of Equivalence, and the definition of \exists

ILLUSTRATION: Not everyone votes votes if and only if someone abstains.

Note that 2119 and 2120 are generalized forms of DeMorgan's Laws.

2121. $\vdash \forall x \forall y A \equiv \forall y \forall x A$

Proof:

$$(1) \forall x \forall y A \vdash A \quad \text{VI (2111)}$$

$$(2) \forall x \forall y A \vdash \forall y \forall x A \quad \text{Gen: .1}$$

$$(3) \vdash \forall x \forall y A \supset \forall y \forall x A \quad \text{Deduction Theorem}$$

$$(4) \vdash \forall y \forall x A \supset \forall x \forall y A \quad \text{Theorem Schema 2121.3}$$

Then use Rule P. ■

ILLUSTRATION:

$$[Raa \wedge Rab \wedge Rac] \wedge [Rba \wedge Rbb \wedge Rbc] \wedge [Rca \wedge Rcb \wedge Rcc] \equiv \\ \vdash [Raa \wedge Rba \wedge Rca] \wedge [Rab \wedge Rbb \wedge Rcb] \wedge [Rac \wedge Rbc \wedge Rcc]$$

2122. $\vdash \exists x \exists y A \equiv \exists y \exists x A$ by 2121, Rule P, and 2117

2123. $\vdash (S_t^x A) \supset \exists x A$ provided that t is a term free for x in A .

Proof:

$$(1) \vdash \forall x \sim A \supset \sim S_t^x A \quad \text{Axiom 4}$$

$$(2) \vdash S_t^x A \supset \sim \forall x \sim A \quad \text{Rule P: .1}$$

■

EXAMPLE: $\vdash P^3xxf^1x \supset \exists yP^3yxf^1y$.

2124. $\vdash \mathbf{A} \supset \exists x\mathbf{A}$

2123

2125. $\vdash \forall x\mathbf{A} \supset \exists x\mathbf{A}$

Rule P: 2102, 2124

2126. Rule of Existential Generalization (\exists Gen). If $\mathcal{H} \vdash S_t^x \mathbf{A}$, where t is a term free for x in \mathbf{A} , then $\mathcal{H} \vdash \exists x\mathbf{A}$.

Proof: By 2123 and MP. ■

EXAMPLE: $\vdash \forall xPxy \supset Pyy$ by Axiom 4, so $\vdash \exists z \cdot \forall xPxz \supset Pzy$ by 2126.

EXAMPLE: If $\mathcal{H} \vdash 3 + 2 = 2 + 3$, then $\mathcal{H} \vdash \exists x \cdot 3 + x = 2 + 3$

REMARK. To illustrate the need for the condition that t be free for x in \mathbf{A} , let \mathbf{A} be $\forall y y < x$, x be x , and t be $y + 1$. Then $S_t^x \mathbf{A}$ is $\forall y y < y + 1$, and $\exists x\mathbf{A}$ is $\exists x \forall y y < x$. The latter does not follow from the former by Existential Generalization, since $y + 1$ is not free for x in $\forall y y < x$.

REMARK. We cannot infer $\exists yPyyx$ from $Pxyx$ by Existential Generalization. For example, we cannot infer $\exists y [y + y = 5]$ (which is a false statement about integers) from $5 + y = 5$ (which is a true statement about integers when y is assigned the value 0).

2127. $\vdash \forall x\mathbf{A} \equiv \mathbf{A}$ if x is not free in \mathbf{A} .

Proof:

- (.1) $\vdash \mathbf{A} \supset \mathbf{A}$
- (.2) $\vdash \mathbf{A} \supset \forall x\mathbf{A}$
- (.3) $\vdash \forall x\mathbf{A} \equiv \mathbf{A}$

Rule P
 $\supset \forall$ Rule (2103): .1
 Rule P: .2, 2102

2128. $\vdash \exists x\mathbf{A} \equiv \mathbf{A}$ if x is not free in \mathbf{A} .

Proof: By 2127, Rule P, and the definition of \exists . ■

2129. $\vdash \forall x[\mathbf{A} \wedge \mathbf{B}] \equiv \forall x\mathbf{A} \wedge \forall x\mathbf{B}$

Proof:

- (.1) $\forall x[\mathbf{A} \wedge \mathbf{B}] \vdash \mathbf{A} \wedge \mathbf{B}$

$\forall I$ (2111): hyp

(.2) $\forall x [A \wedge B] \vdash A$	Rule P: .1
(.3) $\forall x [A \wedge B] \vdash \forall x A$	Gen: .2
(.4) $\forall x [A \wedge B] \vdash B$	Rule P: .1
(.5) $\forall x [A \wedge B] \vdash \forall x B$	Gen: .4
(.6) $\forall x [A \wedge B] \vdash \forall x A \wedge \forall x B$	Rule P: .3, .5
(.7) $\vdash \forall x [A \wedge B] \supset \neg \forall x A \wedge \forall x B$	Deduction Theorem (2116): .6
(.8) $\forall x A \wedge \forall x B \vdash \forall x A$	Rule P: hyp
(.9) $\forall x A \wedge \forall x B \vdash A$	$\forall I$: .8
(.10) $\forall x A \wedge \forall x B \vdash \forall x B$	Rule P: hyp
(.11) $\forall x A \wedge \forall x B \vdash B$	$\forall I$: .10
(.12) $\forall x A \wedge \forall x B \vdash A \wedge B$	Rule P: .9, .11
(.13) $\forall x A \wedge \forall x B \vdash \forall x [A \wedge B]$	Gen: .12
(.14) $\vdash \forall x A \wedge \forall x B \supset \forall x (A \wedge B)$	Deduction Theorem: .13

Then apply Rule P to .7 and .14. ■

ILLUSTRATION:

$$\begin{aligned} & [Pa \wedge Qa] \wedge [Pb \wedge Qb] \wedge [Pc \wedge Qc] \\ & \equiv \neg [Pa \wedge Pb \wedge Pc] \wedge [Qa \wedge Qb \wedge Qc]. \end{aligned}$$

2130. $\vdash \exists x [A \vee B] \equiv \neg \exists x A \vee \exists x B$

Proof:

(.1) $\vdash \forall x [\sim A \wedge \sim B] \equiv \neg \forall x \sim A \wedge \forall x \sim B$	2129
(.2) $\vdash \sim A \wedge \sim B \equiv \sim \neg A \vee \neg B$	Rule P
(.3) $\vdash \forall x \sim [A \vee B] \equiv \neg \forall x \sim A \wedge \forall x \sim B$	2117: .1, .2
(.4) $\vdash \exists x [A \vee B] \equiv \neg \exists x A \vee \exists x B$	Rule P, def of \exists : .3

REMARK. 2129 and 2130 show that a universal quantifier can be distributed over a conjunction, and an existential quantifier can be distributed over a disjunction. ■

2131. $\vdash \exists x [A \vee B] \equiv \neg A \vee \exists x B$ if x is not free in A .
 $\vdash \exists x [B \vee A] \equiv \neg B \vee \exists x A$ if x is not free in B .

Proof: by 2130, 2128, and Rule P. ■

2132. $\vdash \forall x A \vee \forall x B \supset \forall x \bullet A \vee B$.

Proof:

- (.1) $\vdash \forall x A \supset A$ 2102
- (.2) $\vdash \forall x B \supset B$ 2102
- (.3) $\vdash \forall x A \vee \forall x B \supset \bullet A \vee B$ Rule P: .1, .2
- (.4) $\vdash \forall x A \vee \forall x B \supset \forall x \bullet A \vee B$ 2103: .3

■

ILLUSTRATION:

$$\begin{aligned} [Pa \wedge Pb \wedge Pc] \vee [Qa \wedge Qb \wedge Qc] \\ \supset \bullet [Pa \vee Qa] \wedge [Pb \vee Qb] \wedge [Pc \vee Qc]. \end{aligned}$$

Note that the converse need not be true.

2133. $\vdash \forall x [A \vee B] \equiv \bullet A \vee \forall x B$ if x is not free in A .
 $\vdash \forall x [B \vee A] \equiv \bullet \forall x B \vee A$ if x is not free in A .

Proof:

- (.1) $\vdash A \vee \forall x B \supset \forall x \bullet A \vee B$ Rule P: 2127, 2132
- (.2) $\vdash \forall x [A \vee B] \equiv \bullet A \vee \forall x B$ Rule P: .1, Axiom 5
- (.3) $\vdash \forall x [B \vee A] \equiv \bullet \forall x B \vee A$ Rule P, 2117

■

2134. $\vdash \forall x [B \supset A] \equiv \bullet \exists x B \supset A$ if x is not free in A .

Proof:

- (.1) $\vdash \forall x [\sim B \vee A] \equiv \bullet \forall x \sim B \vee A$ 2133
- (.2) $\vdash \forall x [B \supset A] \equiv \bullet \exists x B \supset A$ Rule P: .1

■

2135. Existential Rule (\exists Rule). If $\mathcal{H}, B \vdash A$ and if x is not free in A or in \mathcal{H} , then $\mathcal{H}, \exists x B \vdash A$.

Proof:

- (.1) $\mathcal{H}, B \vdash A$ given
- (.2) $\mathcal{H} \vdash B \supset A$ Deduction Theorem: .1
- (.3) $\mathcal{H} \vdash \forall x \bullet B \supset A$ Gen: .2
- (.4) $\mathcal{H} \vdash \exists x B \supset A$ Rule P: 2134, .3

- (.5) $\mathcal{H}, \exists xB \vdash \exists xB \supset A$
 (.6) $\mathcal{H}, \exists xB \vdash A$

2113: .4
 MP: hyp, .5

■

Frequently in a mathematical argument one establishes (under appropriate assumptions) that there is an element with some property, and then one says “choose such an element, and let it be called y ”. One then continues the argument, assuming that y has the property in question. The following derived rule of inference (not to be confused with the Axiom of Choice, which is much stronger and cannot really be expressed in first-order logic unless set-theoretic postulates are assumed) provides a way of formalizing such an argument.

2136. Rule C.⁴ If $\mathcal{H} \vdash \exists xB$ and $\mathcal{H}, S_y^x B \vdash A$, where y is an individual variable which is free for x in B and which is not free in \mathcal{H} , $\exists xB$, or in A , then $\mathcal{H} \vdash A$. (Note that y may be x .)

Proof:

- (.1) $\mathcal{H}, S_y^x B \vdash A$ given
 (.2) $\mathcal{H}, \exists y S_y^x B \vdash A$ 2135: .1
 (.3) $\mathcal{H} \vdash \exists y S_y^x B \supset A$ Deduction Theorem: .2
 (.4) $\mathcal{H} \vdash \exists xB \supset A$ α Rule (unless y is x)
 (.5) $\mathcal{H} \vdash \exists xB$ given
 (.6) $\mathcal{H} \vdash A$ MP: .4, .5

■

In a proof where $\exists xB$ has been established, the next step is often the introduction of $S_y^x B$ as a new hypothesis, which will eventually be eliminated by Rule C. To make the proof as easy to comprehend as possible, we shall often provide the explanation “choose y ” or “ $\exists I$ ” (existential instantiation) for the new hypothesis, rather than simply indicating that it is a hypothesis.

To illustrate the ideas mentioned above, we consider the problem of proving $\forall x \forall z \forall w \cdot z + x = w + x \supset z = w$ in the theory \mathcal{G} mentioned near the beginning of this section. One can start by assuming $z + x = w + x$ as hypothesis. With the help of postulate (G6) one can prove $\exists y \cdot x + y = 0$, so one chooses such a y by assuming (*) $x + y = 0$ as an additional hypothesis. With the help of (*) and various postulates of \mathcal{G} one can derive

⁴This rule differs from the Rule C of [Rosser 1953], although both rules provide ways of handling acts of choice in an argument.

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

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

$$z + 0 = w + 0$$

$$z = w$$

(Exercise X2105 asks the reader to fill in the details of this proof.) The conclusion $z = w$ says nothing about y , and depends only on the hypothesis $z + x = w + x$, so at this stage hypothesis (*) can be dropped. Of course, in informal mathematics one keeps track of one's hypotheses in an informal way, and just forgets about those that are no longer relevant. In formal logic, however, one needs an explicit rule to eliminate a hypothesis, and this is precisely the purpose of Rule C.

One often finds that the same wff occurs in several lines of a proof, especially if it occurs as a hypothesis. Therefore, for the sake of brevity, it is often convenient to use the number of a line as an abbreviation for the wff which is asserted in that line, when this wff occurs elsewhere in the proof.

Henceforth when a line of a proof is inferred from the immediately preceding line, we shall often omit explicit mention of the preceding line from our justification of the inferred line.

These conventions will be illustrated in the proofs of the next two theorems, which provide examples of applications of Rule C.

2137. $\vdash \exists x \forall z A \supset \forall z \exists x A$.

Proof: If x and z are the same variable, then this wff is equivalent via 2128, 2127, and Rule P to an instance of 2125. Therefore we need consider only the case where x and z are distinct.

- | | | |
|-----|--|-------------------|
| (1) | .1 $\vdash \exists x \forall z A$ | hyp |
| (2) | .1, .2 $\vdash \forall z A$ | choose x |
| (3) | .1, .2 $\vdash A$ | $\forall I$ |
| (4) | .1, .2 $\vdash \exists x A$ | \exists Gen |
| (5) | .1, .2 $\vdash \forall z \exists x A$ | Gen |
| (6) | .1 $\vdash .5$ | Rule C: .1, .5 |
| (7) | $\vdash \exists x \forall z A \supset \forall z \exists x A$ | Deduction Theorem |

■

ILLUSTRATION:

$$\begin{aligned} & [[Raa \wedge Rab \wedge Rac] \vee [Rba \wedge Rbb \wedge Rbc] \vee [Rca \wedge Rcb \wedge Rcc]] \\ & \supset \blacksquare [Raa \vee Rba \vee Rca] \wedge [Rab \vee Rbb \vee Rcb] \wedge [Rac \vee Rbc \vee Rcc] \end{aligned}$$

Note that the converse need not be true.

$$\mathbf{2138.} \vdash \exists x [A \wedge B] \supset \blacksquare \exists x A \wedge \exists x B$$

Proof:

(.1)	.1 $\vdash \exists x .2$	hyp
(.2)	.1, .2 $\vdash A \wedge B$	choose x
(.3)	.1, .2 $\vdash A$	Rule P
(.4)	.1, .2 $\vdash \exists x A$	\exists Gen
(.5)	.1, .2 $\vdash B$	Rule P: .2
(.6)	.1, .2 $\vdash \exists x B$	\exists Gen
(.7)	.1, .2 $\vdash \exists x A \wedge \exists x B$	Rule P: .4, .6
(.8)	.1 $\vdash .7$	Rule C: .1, .7
(.9)	$\vdash \exists x [A \wedge B] \supset \blacksquare \exists x A \wedge \exists x B$	2116

■

ILLUSTRATION:

$$\begin{aligned} & [[P_a \wedge Q_a] \vee [P_b \wedge Q_b] \vee [P_c \wedge Q_c]] \\ & \supset \blacksquare [P_a \vee P_b \vee P_c] \wedge [Q_a \vee Q_b \vee Q_c]. \end{aligned}$$

Note that the converse need not be true.

REMARK. Theorem 2138 is essentially the contrapositive of Theorem 2132 and can be derived from 2132 in a manner analogous to the way 2130 was derived from 2129.

Consider the following attempt at a proof:

(1)	1 $\vdash \exists y Pxy$	hyp
(2)	1, 2 $\vdash Pxx$	choose x
(3)	1, 2 $\vdash \exists v Pv v$	\exists Gen
(4)	1 $\vdash 3$	Rule C: 1, 3 (incorrect)
(5)	$\vdash \exists y Pxy \supset \exists v Pv v$	Deduction Theorem

From an intuitive point of view, the essential error is made in line (2), where the element y , whose existence is asserted in (1), is unjustifiably identified with x . However, one can assume anything one likes as a hypothesis, so from a formal point of view line (2) is correct (though useless). The difficulty arises in line (4), when an attempt is made to eliminate (2) from the list of hypotheses. Note how the restrictions in the statement of Rule C block this incorrect inference. We certainly do not wish to be able to derive (5), since the same reasoning would enable us to infer that $\exists y [8 < y] \supset \exists v [v < v]$ is a true statement about real numbers.

Sometimes it is suggested in connection with Rule C that after deriving $\exists x B$ from a set H of hypotheses, one should simply assert $\exists_y^x B$ (where y is a new variable), but not list it as a hypothesis, since it will eventually be deleted from the list of hypotheses via Rule C. Unfortunately, this can lead to a variety of incorrect inferences, such as the following three:

EXAMPLE A:

- | | |
|---|-------------------|
| (1) .1 ⊢ $\exists x A$ | hyp |
| (2) .1 ⊢ A | choose x |
| (3) .1 ⊢ $\forall x A$ | Gen |
| (4) ⊢ $\exists x A \supset \forall x A$ | Deduction Theorem |

EXAMPLE B:

- | | |
|---|-------------------|
| (1) .1 ⊢ $\forall z \exists x A$ | hyp |
| (2) .1 ⊢ $\exists x A$ | $\forall I$ |
| (3) .1 ⊢ A | choose x |
| (4) .1 ⊢ $\forall z A$ | Gen |
| (5) .1 ⊢ $\exists x \forall z A$ | \exists Gen |
| (6) ⊢ $\forall z \exists x A \supset \exists x \forall z A$ | Deduction Theorem |

EXAMPLE C:

- | | |
|---|----------------|
| (1) .1 ⊢ $\exists x P x \wedge \exists x Q x$ | hyp |
| (2) .1 ⊢ $\exists x P x$ | Rule P |
| (3) .1 ⊢ $P x$ | Choose x |
| (4) .1 ⊢ $\exists x Q x$ | Rule P: .1 |
| (5) .1 ⊢ $Q y$ | choose y |
| (6) .1 ⊢ $P x \wedge Q y$ | Rule P: .3, .5 |
| (7) .1 ⊢ $P x \wedge Q x$ | Sub (2118) |
| (8) .1 ⊢ $\exists x \bullet P x \wedge Q x$ | \exists Gen |
| (9) ⊢ $\exists x P x \wedge \exists x Q x \supset \exists x \bullet P x \wedge Q x$ | 2116 |

Thus one would be led to believe that the converses of 2125, 2137, and 2138 are provable. Actually, we shall see in §23 that they are not. We leave it to the reader to analyze the errors in the examples above.

The proofs of the next three metatheorems require only the Deduction Theorem and Rule P.

2139 Rule of Cases. If $\mathcal{H} \vdash A \vee B$ and $\mathcal{H}, A \vdash C$ and $\mathcal{H}, B \vdash C$, then $\mathcal{H} \vdash C$.

2140 Corollary. If $\mathcal{H}, A \vdash C$ and $\mathcal{H}, \sim A \vdash C$ then $\mathcal{H} \vdash C$.

2141 Indirect Proof (IP). If $\mathcal{H}, \sim A \vdash B$, and $\mathcal{H}, \sim A \vdash \sim B$, then $\mathcal{H} \vdash A$. Also, if $\mathcal{H}, \sim A \vdash B \wedge \sim B$, then $\mathcal{H} \vdash A$.

In Figure 2.2 we provide some rather obvious advice about general strategies which may be used in proving theorems of first-order logic. Of course these strategies may not always provide the most efficient proof of a wff, but in the absence of a better idea they can usually be made to work. When other methods fail to produce a proof, the method of Indirect Proof (2141) can always be forced to work. (See §32.)

We remark that negations can be pushed in via De Morgan's Laws, 2119, and 2120, so they are not considered in Figure 2.2. Of the three strategies presented for proving an implication, the last one (proof by contradiction) is most general. For if $A \vdash B$, then $A, \sim B \vdash B \wedge \sim B$. Similarly, if $\sim B \vdash \sim A$, then $A, \sim B \vdash A \wedge \sim A$. Moreover, if one takes both A and $\sim B$ as hypotheses, one has "more to work with" than if one assumes only one of these wffs.

It will be observed that ingenuity is required in the application of the rules above primarily in the cases of Universal Instantiation and Existential Generalization. Actually, the fundamental problems involved in applying these two rules appropriately are the same. For if one wishes to prove $\exists x A$, instead of proving $S_t^x A$ and applying Existential Generalization, one may assume $\sim \exists x A$ as hypothesis, obtain $\forall x \sim A$ by 2119, then $\sim S_t^x A$ by Universal Instantiation, and derive a contradiction. There is also a fundamental similarity between existential instantiation and Universal Generalization. If one wishes to prove $\forall x A$, instead of proving A and applying Universal Generalization, one may assume $\sim \forall x A$ as hypothesis, obtain $\exists x \sim A$ by 2120, then $\sim A$ by existential instantiation, and derive a contradiction. In either case, one must assure (by an alphabetic change of bound variable, if necessary) that the variable x is not free in any hypothesis.

<u>To Prove</u>	<u>Do</u>
$A \supset B$	Assume A , prove B , and use the Deduction Theorem.
$A \supset B$	Assume $\sim B$, and use the Deduction Theorem and Contrapositive Law $\sim B \supset \sim A \supset \neg A \supset B$.
$A \supset B$	Assume A and $\sim B$, prove $C \wedge \neg C$ (for any wff C), and use the Deduction Theorem and Rule P.
$A \vee B$	Write this as $\sim A \supset B$, and use the strategies for \supset .
$A \wedge B$	Prove A and B separately, then use Rule P.
$\forall x A$	Prove A , and use Gen. (If x is free in a hypothesis, prove $S^x_{\cdot y} A$, for some new variable y , then use Gen and the α Rule.)
$\exists x A$	Prove $S^x_t A$ for some term t , then use \exists Gen. (Of course, the crucial problem is to find a term t such that this is feasible. For complex cases, see the discussion in §30.)

<u>If one has</u>	<u>inferred</u>	<u>Do</u>
$A \wedge B$	Infer A and B by Rule P.	
$A \vee B$	Take A and B as hypotheses separately, and use the Rule of Cases.	
$A \supset B$	Write this as $\sim A \vee B$, and use the strategy for \vee . If A has been established, use Modus Ponens.	
$\forall x A$	Infer $S^x_t A$ for an <i>appropriate</i> term t by $\forall I$. This may have to be done with a number of different terms.	
$\exists x A$	Existentially instantiate by choosing a new variable y (which may be x if it has not already occurred free), and taking $S^x_{\cdot y} A$ as hypothesis. Eliminate the new hypothesis later by Rule C.	

Figure 2.2: General advice on methods of proof

Sometimes attempts to complete a proof appear to be blocked by the restrictions on Rule C. Suppose one has the following fragment of an attempt to prove $\mathcal{H} \vdash A$:

$$\begin{aligned}\mathcal{H}, \sim A &\vdash \exists x B \\ \mathcal{H}, \sim A, B &\vdash B \\ \mathcal{H}, \sim A, B &\vdash \sim B\end{aligned}$$

choose x

It appears that one has now reached the desired contradiction. However, hypothesis B must be eliminated, and Rule C cannot be applied if x occurs free in B . (We may assume that x was chosen as a variable distinct from the free variables in \mathcal{H} and A , but it normally will occur free in B .) Nevertheless, one can complete this proof as follows:

$$\begin{aligned}\mathcal{H}, \sim A, B &\vdash p \wedge \sim p && \text{Rule P} \\ \mathcal{H}, \sim A &\vdash p \wedge \sim p && \text{Rule C} \\ \mathcal{H} &\vdash A && \text{IP}\end{aligned}$$

When one is constructing proofs, it is often convenient to work down from the top and up from the bottom. For example, if one wishes to prove $A \supset B$, one may start by constructing the following proof outline:

$$\begin{aligned}(1) \quad 1 &\vdash A && \text{hyp} \\ (99) \quad 1 &\vdash B \\ (100) \quad \vdash A &\supset B && \text{Ded: 99}\end{aligned}$$

One must then fill in some lines between (1) and (99) so as to justify (99). If B has the form $\forall x C$, and x does not occur free in A , one might expand this outline as follows:

$$\begin{aligned}(1) \quad 1 &\vdash A && \text{hyp} \\ (98) \quad 1 &\vdash C \\ (99) \quad 1 &\vdash \forall x C && \text{Gen: 98} \\ (100) \quad \vdash A &\supset B && \text{Ded: 99}\end{aligned}$$

Working in this way tends to produce well structured proofs.

More formal discussion of proof procedures will be found in Chapter 3. In particular, the reader will find in §30 a summary of particularly useful derived rules of inference.

EXERCISES

X2100. Show that there is an effective test for determining of any wff of \mathcal{F} whether or not it is tautologous.

X2101. In each case below, decide whether the given inference is a correct application of Rule P.

- (a) From $\forall x[Gx \supset Hx] \equiv \exists xHx$ and $\forall x[Gx \supset Hx] \equiv \sim \blacksquare \forall xGx \supset \exists xHx$ infer $[\exists xHx \supset \sim \forall xGx]$.
- (b) From $[q \supset \forall zRz]$ and $[\forall xPx \supset \blacksquare \sim \forall zRz \wedge q]$ infer $\sim \forall xPx$.
- (c) From $[p \supset \blacksquare \forall xQx \vee Ry]$ and $[\forall xQx \supset \blacksquare \exists x[Rx \wedge Qx] \supset Ry]$ and $[\exists x[Rx \wedge Qx] \vee \sim p]$ infer $[p \supset Ry]$.
- (d) From $[[q \wedge r \wedge p] \supset \blacksquare s \supset \forall xPx]$ infer $[p \supset \blacksquare q \supset \blacksquare [r \supset s] \supset \forall xPx]$.

In exercises X2102–X2105, prove the given theorems of \mathcal{G} , the elementary theory of abelian groups which was presented near the beginning of this section.

X2102. $\forall x\forall y \blacksquare x = y \supset y = x$.

X2103. $\forall x\forall y \blacksquare x = y \supset \blacksquare z = x \supset z = y$.

X2104. $\forall x\forall y\forall z \blacksquare x = y \wedge y = z \supset x = z$.

X2105. $\forall x\forall z\forall w \blacksquare z + x = w + x \supset z = w$.

Prove the following wffs, using the derived rules of inference discussed in this section. Present your proofs in the manner of the proofs of theorem schemata 2137 and 2138.⁵

X2106. $\forall x[Rx \supset Px] \wedge \forall x[\sim Qx \supset Rx] \supset \forall x \blacksquare Px \vee Qx$.

X2107. $Rab \wedge \forall x\forall y[Rxy \supset \blacksquare Ryx \wedge Qxy] \wedge \forall u\forall v[Quv \supset Quu] \supset \blacksquare Qaa \wedge Qbb$.

X2108. $\forall x\exists y \blacksquare Px \supset Py$.

X2109. $\exists x[p \wedge Qx] \equiv \blacksquare p \wedge \exists xQx$.

X2110. $\exists xRx \wedge \forall y[Ry \supset \exists zQyz] \wedge \forall x\forall y[Qxy \supset Qxx] \supset \exists x\exists y \blacksquare Qxy \wedge Ry$.

⁵See the end of the Preface for information about the ETPS program, which facilitates work on exercises such as these.

X2111. $\forall x [\exists y Pxy \supset \forall y Qxy] \wedge \forall z \exists y Pzy \supset \forall y \forall x Qxy.$

X2112. $\exists v \forall x Pxv \wedge \forall x [Sx \supset \exists y Qyx] \wedge \forall x \forall y [Pxy \supset \sim Qxy]$
 $\supset \exists u \sim Su.$

X2113. $\forall y \exists w Ryw \wedge \exists z \forall x [Px \supset \sim Rzx] \supset \exists x \sim Px.$

X2114. $\forall x Rx b \wedge \forall y [\exists z Ryz \supset Ray] \supset \exists u \forall v Ruv.$

X2115. $\forall x [\exists y Pxy \supset \forall z Pzz] \wedge \forall u \exists v [Puv \vee \neg Mu \wedge Qf^2uv]$
 $\wedge \forall w [Qw \supset \sim Mg^1w] \supset \forall u \exists v \neg Pg^1uv \wedge Pvu.$

X2116. $\forall x \exists y [Px \supset \neg Rxy^1h^1y \wedge Py] \wedge \forall w [Pw \supset \neg Pg^1w \wedge Ph^1w]$
 $\supset \forall x \neg Px \supset \exists y \neg Rxy \wedge Py.$

X2117. $\forall u \forall v [Ruu \equiv Ruv] \wedge \forall w \forall z [Rww \equiv Rzw] \supset \neg \exists x Rxx \supset \forall y Ryy.$

X2118. $\forall x [[p \wedge Qx] \vee \neg p \wedge Rx] \supset \neg \forall x Qx \vee \forall x Rx.$

Hint: If it is true every day that the roof was fixed and the rug stays dry, or that the roof was not fixed and the rug gets wet, then the rug stays dry every day or the rug gets wet every day.

X2119. $\exists y \forall x \neg Py \supset Px.$

X2120. $\forall u \forall v \forall w [Puv \vee Pvw] \supset \exists x \forall y Pxy.$

X2121. $\exists v \forall y \exists z \neg [Payh^1y \vee Pvyyf^1y] \supset Pvyz.$

X2122. $[\exists x Rxx \supset \forall y Ryy] \supset \exists u \forall v \neg Ruu \supset Rvv.$

X2123. $\exists y [Py \supset Qx] \supset \exists y \neg Py \supset Qy.$

X2124. $\exists x [Px \supset Qx] \equiv \neg \forall x Px \supset \exists x Qx.$

X2125. $\exists x \forall y [Px \equiv Py] \equiv \neg \exists x Px \equiv \forall y Py.$

X2126. $\forall x [Px \equiv \exists y Py] \equiv \neg \forall x Px \equiv \exists y Py.$

X2127. $\exists x \forall y [Py \equiv Px] \supset \neg \forall x Px \vee \forall x \sim Px.$

X2128. $\forall x [Px \equiv \forall y Py] \equiv \neg \exists x Px \equiv \forall y Py.$

X2129. $[\exists x \forall y [Px \equiv Py] \equiv \neg \exists x Qx \equiv \forall y Py]$
 $\equiv \neg \exists x \forall y [Qx \equiv Qy] \equiv \neg \exists x Px \equiv \forall y Qy.$

X2130. $\forall x Px \supset \neg \exists y Qy \vee \exists z \neg Pz \supset Qz.$

X2131. $\forall x Px \supset \exists y \neg \forall x \forall z Qxyz \supset \sim \forall z \neg Pz \wedge \sim Qyyz.$

X2132. $\forall w \sim Rww \supset \exists x \exists y \bullet \sim Rxy \wedge \bullet Qyx \supset \forall z Qzz.$

X2133. $\forall x [\exists y Qxy \supset Px] \wedge \forall v \exists u Quv \wedge \forall w \forall z [Qwz \supset \bullet Qzw \vee Qzz] \supset \forall z Pz.$

X2134. $\forall z \exists x [\forall y Pxy \vee Qxz] \supset \forall y \exists x [Pxy \vee Qxy].$

X2135. $\exists x \forall y [Px \wedge Qy \supset \bullet Qx \vee Py].$

X2136. $\exists x \exists y \forall u [Pxyz \supset Puxx].$

X2137. $\exists x \forall y [Px \supset \bullet Qx \vee Py].$

X2138. $\forall x \exists y Fxy \wedge \exists x \forall \epsilon \exists n \forall w [Snw \supset Dwx\epsilon] \wedge \forall \epsilon \exists \delta \forall x_1 \forall x_2 [Dx_1x_2\delta \supset \forall y_1 \forall y_2 \bullet Fx_1y_1 \wedge Fx_2y_2 \supset Dy_1y_2\epsilon] \supset \exists y \forall \epsilon \exists m \forall w \bullet Smw \supset \forall z \bullet Fwz \supset Dzy\epsilon.$

(Here δ , ϵ , n , and m are used as individual variables.)

X2139. $(\exists x \epsilon S) Px \equiv \sim (\forall x \epsilon S) \sim Px.$

X2140. $(\forall x \epsilon S) Px \equiv \sim (\exists x \epsilon S) \sim Px.$

X2141. Can $\exists x \forall y Pyx$ be inferred from $\forall y Py fy$ by the rule of Existential Generalization? (Here f is a monadic function constant.) Explain exactly how this inference does or does not conform to this rule.

In exercises X2142–X2143, you are given alleged proofs to evaluate. Decide whether each line of the proof is correctly inferred from the specified line(s) by the specified rule of inference. Treat each line of the proof separately; when dealing with one line, you need not be concerned with the question of whether other lines were obtained correctly.

X2142.

- | | | |
|------|---|--------------|
| (1) | $1 \vdash Px \wedge \forall x \exists y Rxyz$ | hyp |
| (2) | $1 \vdash Px$ | Rule P: 1 |
| (3) | $1 \vdash Pu$ | Sub: 2 |
| (4) | $1 \vdash \exists y Ruyz$ | UI: 1 |
| (5) | $1 \vdash \forall z \exists y Ruyz$ | UGen: 4 |
| (6) | $1 \vdash \exists y Ruyu$ | UI: 5 |
| (7) | $1, 7 \vdash Ruzu$ | hyp |
| (8) | $1, 7 \vdash \exists v Rvvu$ | EGen: 7 |
| (9) | $1 \vdash \exists v Rvvu$ | Rule C: 6, 8 |
| (10) | $1, 7 \vdash \exists x Rxzu$ | EGen: 7 |

- (11) $1 \vdash \forall y \sim Ruyz \supset \sim Ruuu$ Rule P: 4
 (12) $1 \vdash \exists y Ruyy$ UI: 5

X2143.

- (1) $1 \vdash \forall x [Px \supset \sim Qxy \vee Rxy] \supset \sim Rxy \supset Qxy$ hyp
- (2) $1 \vdash [Px \supset \sim Qxy \vee Rxy] \supset \sim Rxy \supset Qxy$ $\forall I$: 1
- (3) $1 \vdash \sim Qxy \supset Rxy$ Rule P: 2
- (4) $1 \vdash \forall y \sim Qxy \supset Rxy$ Gen: 3
- (5) $1 \vdash \exists z \forall y \sim Qxy \supset Rzy$ \exists Gen: 4
- (6) $1, 6 \vdash \forall y \sim Qxy \supset Rzy$ choose z
- (7) $1, 6 \vdash \sim Qxy \supset Rzx$ $\forall I$: 6
- (8) $1, 6 \vdash \exists x \sim Qxx \supset Rzx$ \exists Gen: 7
- (9) $1 \vdash \exists x \sim Qxx \supset Rzx$ Rule C: 5, 8
- (10) $\vdash \forall x [Px \supset \sim Qxy \vee Rxy] \supset \sim Rxy \supset Qxy \supset \exists x \sim Qxx \supset Rzx$ Deduction Theorem: 9

X2144. Let \mathcal{F}^* be the logistic system obtained by adding to the system \mathcal{F} the additional axiom schema $\exists x A \supset \forall x A$. Prove that \mathcal{F}^* is absolutely consistent.

X2145. Let \mathcal{F}^{**} be the logistic system obtained by adding to the system \mathcal{F} the additional axiom schema $\forall x \exists y A \supset \exists y \forall x A$. Is \mathcal{F}^{**} is absolutely consistent?

X2146. Express the following statements in symbolic form, and prove that the negation of statement A implies statement B.

- (A) Everyone who bought a car also bought a trailer or a boat without a trailer.
- (B) Someone bought a car but no boat and no trailer.

§22. Prenex Normal Form

In this section we wish to show that each wff is equivalent to a wff in which no quantifier is in the scope of a propositional connective. However, we wish to consider existential as well as universal quantifiers, and ignore the negations which occur as parts of existential quantifiers. To be precise about this, we shall introduce a language \mathcal{E} in which existential as well as universal quantifiers are primitive, and formulate certain definitions in terms of wffs of \mathcal{E} .

DEFINITION. The primitive symbols of \mathcal{E} are those of \mathcal{F} , plus the symbol \exists . The formation rules of \mathcal{E} are those of \mathcal{F} , plus the following:

- (g) If B is a wff of \mathcal{E} and x is an individual variable, then $\exists xB$ is a wff of \mathcal{E} .

Clearly each wff of \mathcal{F} is a wff of \mathcal{E} . Also, every wff of \mathcal{E} is an abbreviation for a wff of \mathcal{F} , under the convention that $\exists xB$ is an abbreviation for $\sim \forall x \sim B$. Thus, $\exists xB$ and $\sim \forall x \sim B$ are distinct wffs of \mathcal{E} , but are abbreviations for the same wff of \mathcal{F} . In this section, when we assert of a wff A of \mathcal{E} that $\vdash A$, we shall mean that the wff of \mathcal{F} for which A is an abbreviation is a theorem of \mathcal{F} .

DEFINITION. An occurrence of a quantifier $\forall x$ or $\exists x$ is *vacuous* iff its variable x has no free occurrence in its scope. Thus, if x is not free in B , then the exhibited quantifier in $\forall xB$ is vacuous.

DEFINITION. A wff G of \mathcal{E} is in *prenex normal form* iff no vacuous quantifier occurs in G and no quantifier in G occurs in the scope of a propositional connective of G . A wff A of \mathcal{F} is in prenex normal form iff there is a wff of \mathcal{E} which is in prenex normal form and which is an abbreviation for A . Thus, a wff in prenex normal form is either quantifier-free, or is of the form $\exists_1 x_1 \dots \exists_n x_n M$, where M is a quantifier-free wff and for each i , $\exists_i x_i$ is $\forall x_i$ or $\exists x_i$. Here M is called the *matrix* of the wff, and the sequence $\exists_1 x_1 \dots \exists_n x_n$ of quantifiers is called the *prefix*. If G is a wff in prenex normal form such that $\vdash C \equiv G$, we shall say that G is a *prenex normal form* of C .

DEFINITION. A wff D is *rectified* iff D contains no vacuous quantifier, no variable occurs both bound and free in D , and no two quantifier-occurrences in D bind the same variable. Clearly a wff in prenex normal form must be rectified.

For any rectified wff D of \mathcal{E} , we define the *prenex normal form* of D as follows. Let $(\forall x)'$ be $\exists x$ and $(\exists x)'$ be $\forall x$. Let the quantifiers in D be $\exists_1 x_1, \dots, \exists_n x_n$, in the left-to-right order in which they occur in D (so $n \geq 0$ and each \exists_i is \forall or \exists). We say that $\exists_i x_i$ occurs *positively* [*negatively*] in D iff $\exists_i x_i C_i$, where C_i is the scope of the quantifier $\exists_i x_i$, is a positive [negative] wf part of D . The prenex normal form of D is $Z_1 x_1 \dots Z_n x_n M$, where M is the wff obtained from D by erasing all quantifiers of D , and for $1 \leq i \leq n$, $Z_i x_i$ is $\exists_i x_i$ if $\exists_i x_i$ occurs positively in D , but $Z_i x_i$ is $(\exists_i x_i)'$ if $\exists_i x_i$ occurs negatively in D . Examples will be given after we prove the following fundamental theorem.

2200 Prenex Normal Form Theorem. Every wff has a prenex normal form. If D is a rectified wff and G is the prenex normal form of D , then $\vdash D \equiv G$.

Proof: For any wff D of \mathcal{E} , let $\#D$ be the sum, over all occurrences of quantifiers D , of the number of occurrences of propositional connectives in whose scope lies the given quantifier. We first show by induction on $\#D$ that if D is any rectified wff of \mathcal{E} and G is the prenex normal form of D , then $\vdash D \equiv G$.

If $\#D = 0$, clearly G is D , so $\vdash D \equiv G$ by Rule P.

If $\#D > 0$, some quantifier of D is in the scope of some propositional connective. The left-most such quantifier of D is immediately in the scope of some propositional connective. Let M be the wf part of D consisting of that connective and its scope. (Thus, M is the smallest wf part of D which properly includes the left-most quantifier of D which is in the scope of some propositional connective in D .) For each of the possible forms of M , let N be the wff indicated below:

<u>M</u>	<u>N</u>
$\sim \forall x B$	$\exists x \sim B$
$\sim \exists x B$	$\forall x \sim B$
$A \vee \forall x B$	$\forall x [A \vee B]$
$\forall x B \vee A$	$\forall x [B \vee A]$
$A \vee \exists x B$	$\exists x [A \vee B]$
$\exists x B \vee A$	$\exists x [B \vee A]$

(Note that in each case x is not free in A , since D is rectified.) Let E be the wff obtained when M is replaced by N in D . $\vdash M \equiv N$ by 2120, 2119, 2133, 2131, and Rule P, so $\vdash D \equiv E$ by the Substitutivity of Equivalence. Clearly E is a rectified wff, and the bound variables of E have their quantifiers in the same left-to-right order in E as in D (since if M is $A \vee \forall x B$ or $A \vee \exists x B$, then A must be quantifier-free). Hence it is easy to see that the prenex normal form G of D is also the prenex normal form of E . Also, $\#E = \#D - 1$, so by inductive hypothesis, $\vdash E \equiv G$. Hence $\vdash D \equiv G$ by Rule P. This completes the inductive proof.

Now for any wff C of \mathcal{F} , let D be a rectified wff obtained from C by deleting any vacuous quantifiers and by making any necessary alphabetic changes of bound variables. $\vdash C \equiv D$ by 2127, 2109, and the Substitutivity of Equivalence (2106). Hence if G is the prenex normal form of D , $\vdash C \equiv G$, so C has a prenex normal form. ■

EXAMPLE: Find a prenex normal form of $\forall x \exists y Pxy \equiv \exists z Qzx$.

Eliminating the \equiv sign we obtain

$$[\forall x \exists y Pxy \supset \exists z Qzx] \wedge [\exists z Qzx \supset \forall x \exists y Pxy],$$

which can be rectified to

$$[\forall x_2 \exists y P_{x_2}y \supset \exists z Q_{zx}] \wedge [\exists z_1 Q_{z_1}x \supset \forall x_1 \exists y_1 P_{x_1}y_1].$$

In this wff the quantifiers $\forall x_2$, $\exists y$, and $\exists z_1$ occur negatively, and the others occur positively, so a prenex normal form of this wff is

$$\exists x_2 \forall y \exists z \forall z_1 \forall x_1 \exists y_1 [[P_{x_2}y \supset Q_{zx}] \wedge \neg Q_{z_1}x \supset P_{x_1}y_1].$$

REMARK. When students are confronted with an example like the one above, they are sometimes tempted to change its *free* variables. However, there is no justification for doing this, and the resulting wff is usually not equivalent to the original one. Note Exercise X2312 below.

REMARK. Sometimes by using theorem schemata 2129 and 2130 one can reduce the number of quantifiers in a wff. Let us apply this idea to find an alternative prenex normal form of the wff dealt with in the example above. We leave it to the reader to explain why each wff below is equivalent to the preceding one:

$$\begin{aligned} & [\forall x_2 \exists y P_{x_2}y \supset \exists z Q_{zx}] \wedge [\exists z_1 Q_{z_1}x \supset \forall x_1 \exists y_1 P_{x_1}y_1] \\ & [\exists x_2 \forall y \sim P_{x_2}y \vee \exists z Q_{zx}] \wedge [\forall z_1 \sim Q_{z_1}x \vee \forall x_1 \exists y_1 P_{x_1}y_1] \\ & [\exists z \forall y \sim P_{zy} \vee \exists z Q_{zx}] \wedge [\forall y \sim Q_{yx} \vee \forall x_1 \exists y_1 P_{x_1}y_1] \\ & \exists z [\forall y \sim P_{zy} \vee Q_{zx}] \wedge \forall y \forall x_1 \exists y_1 [\sim Q_{yx} \vee P_{x_1}y_1] \\ & \exists z \neg \forall y [\sim P_{zy} \vee Q_{zx}] \wedge \forall y \forall x_1 \exists y_1 [\sim Q_{yx} \vee P_{x_1}y_1] \\ & \exists z \forall y \neg [P_{zy} \supset Q_{zx}] \wedge \forall x_1 \exists y_1 [Q_{yx} \supset P_{x_1}y_1] \\ & \exists z \forall y \forall x_1 \exists y_1 \neg [P_{zy} \supset Q_{zx}] \wedge \neg Q_{yx} \supset P_{x_1}y_1. \end{aligned}$$

EXERCISES

Find the prenex normal form of each of the rectified wffs below:

X2200. $\sim \forall x R_{xu} \wedge \exists y \neg \sim \forall z R_{yz} \supset Q_{uy}$

X2201. $\sim \neg \exists x Q_{xy} \supset \sim \forall z \neg \forall u P_{uyz} \supset \sim \exists v R_v$

Find prenex normal forms for the wffs below:

X2202. $\forall xPx \equiv \exists y Qy$

X2203. $\forall x \sim \exists y \forall u Pxy \supset \sim \exists z Pxy \supset \forall y Puy$

X2204. $\forall x [Pxu \vee \exists y \exists Qxy \wedge \sim \forall u Ruz] \supset \forall z Qzu$

X2205. $\forall w \exists x [Pwx \supset \forall y Pyx] \supset \forall x Qxw$

X2206. $\sim [\forall u Puz \wedge \sim \exists x \exists Qux \vee \forall z \sim Rxz] \supset \exists z Quz$

X2207. $\forall y [Py \wedge \exists y Qy] \supset \forall x Rxy \supset \exists z Rzx$

X2208. $\forall xPx \supset \exists x \exists z Qxz \supset \exists y \forall x Rxyz$

X2209. $\forall z [\sim \exists x [Px \wedge \forall y Rxyz] \vee \forall z \sim \exists y Qyz \supset Px]$

For each of the wffs below, find a prenex normal form in which the number of occurrences of quantifiers is minimized.

X2210. $\forall u \exists v \sim \exists w \exists x \forall y \forall z [Pwz \vee Qy] \supset \sim Rux$

X2211. $\sim [\sim \forall x [Qx \vee \sim \forall y Ry] \vee \forall z Qz \vee \sim \forall w Rw]$

X2212. $\forall xPx \supset \exists x \exists z Qxz \supset \exists y \forall x Rxyz$

X2213. $\forall z [\sim \exists x [Px \wedge \forall y Rxyz] \vee \forall z \sim \exists y Qyz \supset Px]$

X2214. $\forall xPx \equiv \exists y \forall z Qxyz$

Prove the following wffs:

X2215. $\exists x \forall y \exists z [Pzx \wedge Pxy \supset Pyy]$

X2216. $\exists x \forall y \exists Qx \supset [Px \vee Qy]$

X2217. $\exists x \exists y \forall z [Gxy \supset [Gzx \wedge [Hxy \supset Hz]]]$

§23. Semantics of \mathcal{F}

DEFINITION. An *interpretation* $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$ of \mathcal{F} consists of a non-empty set \mathcal{D} called the *universe* or the *domain of individuals* and (if there are constants in the formulation of \mathcal{F} under consideration) a mapping \mathcal{J} defined on the constants of \mathcal{F} satisfying the conditions below. Given an interpretation \mathcal{M} ,

an *assignment* into \mathcal{M} is a function φ defined on the variables of \mathcal{F} and satisfying the conditions below.

- (a) If x is an individual variable [constant], then $\varphi x [\mathcal{J}x]$ is a member of \mathcal{D} .
- (b) If f^n is an n -ary function variable [constant], then $\varphi f^n [\mathcal{J}f^n]$ is a function from \mathcal{D}^n into \mathcal{D} , where \mathcal{D}^n is the set of ordered n -tuples of elements in \mathcal{D} .
- (c) If p is a propositional variable [constant], then $\varphi p [\mathcal{J}p]$ is a truth value.
- (d) If P^n is an n -ary predicate variable [constant], then $\varphi P^n [\mathcal{J}P^n]$ is a function from \mathcal{D}^n into the set $\{\text{T}, \text{F}\}$ of truth values.

DEFINITIONS. Let φ and ψ be assignments, x be a variable (of any type), and S be a set of variables. We say that φ and ψ *agree off* x iff $\varphi y = \psi y$ for all variables y (of any type) distinct from x . We say that φ and ψ *agree on* the variables in S iff $\varphi y = \psi y$ for all variables y in S .

We shall next define the semantical values (meanings) of certain expressions, so that the values of terms are individuals, and the values of wffs are truth values.

DEFINITION. Given an interpretation $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$ and an assignment φ , we define $\mathcal{V}_\varphi^\mathcal{M} Z$, the *value of Z with respect to φ in \mathcal{M}* , for each variable, constant, term, or wff Z , as follows by induction on the construction of Z :

- (1) $\mathcal{V}_\varphi^\mathcal{M} Z = \varphi Z$ if Z is a variable.
- (2) $\mathcal{V}_\varphi^\mathcal{M} Z = \mathcal{J}Z$ if Z is a constant.
- (3) If F^n is an n -ary function or predicate variable or constant, and t_1, \dots, t_n are terms $\mathcal{V}_\varphi^\mathcal{M}(F^n t_1, \dots, t_n) = (\mathcal{V}_\varphi^\mathcal{M} F^n)(\mathcal{V}_\varphi^\mathcal{M} t_1, \dots, \mathcal{V}_\varphi^\mathcal{M} t_n)$, i.e., the value of the function $\mathcal{V}_\varphi^\mathcal{M} F^n$ on the indicated arguments.
- (4) $\mathcal{V}_\varphi^\mathcal{M} \sim A = \text{T}$ if $\mathcal{V}_\varphi^\mathcal{M} A = \text{T}$.
 $= \text{F}$ if $\mathcal{V}_\varphi^\mathcal{M} A = \text{F}$.
- (5) $\mathcal{V}_\varphi^\mathcal{M} [A \vee B] = \text{T}$ if $\mathcal{V}_\varphi^\mathcal{M} A = \text{T}$ or $\mathcal{V}_\varphi^\mathcal{M} B = \text{T}$.
 $= \text{F}$ if $\mathcal{V}_\varphi^\mathcal{M} A = \text{F}$ and $\mathcal{V}_\varphi^\mathcal{M} B = \text{F}$.
- (6) $\mathcal{V}_\varphi^\mathcal{M} \forall x A = \text{T}$ if $\mathcal{V}_\psi^\mathcal{M} A = \text{T}$ for every assignment ψ which agrees with φ off x .
 $= \text{F}$ if $\mathcal{V}_\psi^\mathcal{M} A = \text{F}$ for some assignment ψ which agrees with φ off x .

EXAMPLE: Let us consider $\mathcal{V}^{\mathcal{M}} \mathcal{M}_{\varphi} \forall x Pyx$, where $\mathcal{M} = \langle \mathbb{N}, \mathcal{J} \rangle$, \mathbb{N} is the set of natural numbers, and for all n and m in \mathbb{N} , $(\mathcal{J}P) nm = T$ iff $n \leq m$ (so that Pyx means $y \leq x$). Suppose $\varphi y = 0$, and let ψ be an assignment which agrees with φ off x . Then $\psi y = \varphi y = 0$, and $\mathcal{V}_{\psi}^{\mathcal{M}} Pyx = (\mathcal{J}P)(\psi y)(\psi x) = (\mathcal{J}P) 0(\psi x) = T$, since $0 \leq \psi x$ (no matter what natural number ψx is). Therefore $\mathcal{V}_{\varphi}^{\mathcal{M}} \forall x Pyx = T$ when $\varphi y = 0$. Next, suppose $\varphi y = 5$. Let ψ be the assignment which agrees with φ off x , with $\psi x = 4$. Then $\mathcal{V}_{\psi}^{\mathcal{M}} Pyx = (\mathcal{J}P)(\psi y)(\psi x) = (\mathcal{J}P) 5 4 = F$ (since it is false that $5 \leq 4$), so $\mathcal{V}_{\varphi}^{\mathcal{M}} \forall x Pyx = F$ in this case.

It is important that we have a rigorous way of defining semantical concepts, but the reader should not let the formalism obscure the basic ideas. The example above can be summarized informally by saying that $\forall x y \leq x$ is true in the natural numbers when y is 0, but false when y is 5. It is obvious that some value must be assigned to the free variable y before one can consider whether the wff is true or false.

REMARK. We shall sometimes write $\mathcal{V}_{\varphi}^{\mathcal{M}}$ as \mathcal{V}_{φ} , $\mathcal{V}^{\mathcal{M}}$, or \mathcal{V} , when the interpretation or assignment is clear from the context, or irrelevant.

For convenience, we shall on occasion use “ \sim ” and “ \vee ” in our meta-language to denote the truth functions negation and disjunction, respectively. The context will reveal when these symbols are being used as symbols of the object language \mathcal{F} , and when they are being used as notations in the meta-language. Thus, (4) might be written $\mathcal{V}_{\varphi} \sim A = \sim \mathcal{V}_{\varphi} A$ and (5) might be written $\mathcal{V}_{\varphi} [A \vee B] = (\mathcal{V}_{\varphi} A) \vee (\mathcal{V}_{\varphi} B)$.

2300 Proposition. Let \mathcal{M} be an interpretation, Z a term or wff, and φ and ψ assignments into \mathcal{M} which agree on all variables which occur free in Z . Then $\mathcal{V}_{\varphi}^{\mathcal{M}} Z = \mathcal{V}_{\psi}^{\mathcal{M}} Z$.

The proof is by induction on the construction of Z . We leave the details to the reader.

DEFINITIONS. Let A be a wff, \mathcal{M} be an interpretation, and φ be an assignment into \mathcal{M} .

- (1) φ satisfies A in \mathcal{M} iff $\mathcal{V}_{\varphi}^{\mathcal{M}} A = T$.
- (2) A is satisfiable in \mathcal{M} iff there is an assignment φ which satisfies A in \mathcal{M} .
- (3) A is satisfiable iff there is an interpretation in which A is satisfiable.

- (4) A set \mathcal{G} of wffs is (*simultaneously*) *satisfiable* iff there is an interpretation \mathcal{M} and an assignment φ such that for all $\mathbf{B} \in \mathcal{G}$, $\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{B} = \top$.
- (5) \mathbf{A} is *valid* in \mathcal{M} iff for every assignment φ into \mathcal{M} , $\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A} = \top$.
- (6) \mathbf{A} is *valid* iff \mathbf{A} is valid in every interpretation.
- (7) A sentence \mathbf{A} is *true* in \mathcal{M} iff $\mathcal{V}^{\mathcal{M}} \mathbf{A} = \top$ and *false* in \mathcal{M} iff $\mathcal{V}^{\mathcal{M}} \mathbf{A} = \perp$.
- (8) \mathbf{A} is *contradictory*, or *unsatisfiable*, iff for every interpretation \mathcal{M} and every assignment φ into \mathcal{M} , $\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A} = \perp$.
- (9) A set of wffs is *unsatisfiable* iff it is not simultaneously satisfiable.
- (10) A *model* for a set of wffs is an interpretation in which each of the wffs is valid. A *model* for a first order theory is a model for the set of postulates of the theory. (Normally we speak of models for sets of *sentences*.)
- (11) A *countermodel* for a sentence is an interpretation in which the sentence is false. (It is a counterexample to the conjecture that the sentence is valid.)

NOTATION. We write $\mathcal{M} \models_\varphi \mathbf{A}$, $\mathcal{M} \models \mathbf{A}$, and $\models \mathbf{A}$ to indicate that φ satisfies \mathbf{A} in \mathcal{M} , \mathbf{A} is valid in \mathcal{M} , and \mathbf{A} is valid, respectively. The notation $\mathcal{M} \models \mathbf{A} [\varphi]$ is often used instead of $\mathcal{M} \models_\varphi \mathbf{A}$.

REMARK. We shall show that the theorems of \mathcal{F} are precisely the valid wffs.

EXAMPLES: We add to \mathcal{F} an individual constant 1, a binary function constant ${}^2+$, and predicate constants E^1 , O^1 , $<^2$, $=^2$. We shall let $(s + t)$, $s < t$, and $s = t$ serve as abbreviations for 2st , $<^2st$, and $=^2st$, respectively. We consider two interpretations, \mathbb{N} and \mathbb{R} . The domain of individuals is the set $\{0, 1, 2, 3, \dots\}$ of natural numbers for \mathbb{N} , and the set of real numbers for \mathbb{R} . In both interpretations the constants 1, $+$, E , O , $<$, $=$ denote the number one, the addition function, the set of even integers, the set of odd integers, and the strict ordering and equality relations, respectively.

- (a) $x + y = 1$ is satisfiable but not valid in \mathbb{N} and in \mathbb{R} .
- (b) $\exists x \exists y [x + y = 1]$ is true in \mathbb{N} and in \mathbb{R} .
- (c) $\forall x \forall y [x + y = 1]$ is false in \mathbb{N} and in \mathbb{R} .
- (d) $\forall x \exists y [x + y = 1]$ is false in \mathbb{N} , true in \mathbb{R} . Thus, it is satisfiable but not valid.
- (e) $\exists y \forall x [x + y = 1]$ is false in \mathbb{N} and in \mathbb{R} .
- (f) $\forall x \exists y [x + y = 1] \supset \exists y \forall x [x + y = 1]$ is true in \mathbb{N} but false in \mathbb{R} , and so is not valid.
- (g) $\exists x O x \supset \forall x O x$ is false in \mathbb{N} , and so is not valid.
- (i) $\forall x \exists y x < y$ is true in \mathbb{N} .

- (j) $\exists y y < y$ is false in \mathbb{N} .
- (k) $\forall x \exists y x < y \supset \exists y y < y$ is false in \mathbb{N} , and so is not valid. This illustrates the need for the restriction on Axiom 4.
- (l) $\forall x [x = y \supset y = x]$ is valid in \mathbb{N} .
- (m) $\mathcal{V}_\varphi^{\mathbb{N}} [x = y \supset \forall x \bullet y = x] = F$ when $\varphi x = \varphi y$, so this wff is not valid in \mathbb{N} .
- (n) $\forall x [x = y \supset y = x] \supset [x = y \supset \forall x \bullet y = x]$ is not valid in \mathbb{N} , hence not valid. This illustrates the need for the restriction on Axiom 5. Note, however, that this wff is satisfied in \mathbb{N} by any assignment φ such that $\varphi x \neq \varphi y$.

2301 Proposition. Let A be a wff and \mathcal{M} be an interpretation.

- (a) A is valid [in \mathcal{M}] iff $\sim A$ is not satisfiable [in \mathcal{M}].
- (b) A is satisfiable [in \mathcal{M}] iff $\sim A$ is not valid [in \mathcal{M}].
- (c) A is valid [in \mathcal{M}] iff $\forall x A$ is valid [in \mathcal{M}].
- (d) A is satisfiable [in \mathcal{M}] iff $\exists x A$ is satisfiable [in \mathcal{M}].

We leave the proofs of these statements to the reader. However, the proof of (a) is essentially conveyed by the following abbreviated argument:

$$\begin{aligned} A \text{ is valid} &\Leftrightarrow \forall \mathcal{M} \forall \varphi \mathcal{M} \models_\varphi A \\ &\Leftrightarrow \sim \exists \mathcal{M} \exists \varphi \sim (\mathcal{M} \models_\varphi A) \\ &\Leftrightarrow \sim \exists \mathcal{M} \exists \varphi \mathcal{M} \models_\varphi \sim A \\ &\Leftrightarrow \sim A \text{ is not satisfiable}. \end{aligned}$$

DEFINITIONS. For any wff B of which the free variables are precisely x_1, \dots, x_n , we define the *universal closure* A of B to be $\forall x_1 \dots \forall x_n B$, and the *existential closure* E of B to be $\exists x_1 \dots \exists x_n B$. Here A and E are closed wffs, and if B is a closed wff, then $A = B = E$. Clearly $\vdash A \supset B$ and $\vdash B \supset E$, and $\vdash B$ iff $\vdash A$. Also, for any interpretation \mathcal{M} , B is valid [in \mathcal{M}] iff A is valid [in \mathcal{M}], and B is satisfiable [in \mathcal{M}] iff E is satisfiable [in \mathcal{M}].

If θ is any proper substitution for free occurrences of variables and A is a wff, we say that θ is *free on* A iff for each free variable m of A , θm is free for m in A . Note that if t is free for x in A , then S_t^x is free on A .

Given an interpretation \mathcal{M} , an assignment φ into \mathcal{M} , and a proper substitution θ for free occurrences of variables, we define $\varphi \circ \theta$ to be that assignment into \mathcal{M} such that for each variable m , $(\varphi \circ \theta)m = \mathcal{V}_\varphi(\theta m)$.

2302 Substitution-Value Theorem. If A is a variable, constant, term, or wff, and θ is a proper substitution for free occurrences of variables such that θ is free on A , and φ is any assignment, then $\mathcal{V}_\varphi \theta A = \mathcal{V}_{\varphi \circ \theta} A$.

REMARK. Given an interpretation and an assignment φ , $\mathcal{V}_\varphi \mathbf{m} = \varphi \mathbf{m}$ for each variable \mathbf{m} , so \mathcal{V}_φ is an extension of φ . Therefore, if \mathbf{K} is a constant, term, or wff, we could write $\varphi \mathbf{K}$ as an abbreviation for $\mathcal{V}_\varphi \mathbf{K}$. In this notation the Substitution-Value Theorem says that $\varphi(\theta \mathbf{A}) = (\varphi \circ \theta) \mathbf{A}$.

Corollary. If \mathbf{t} is a term free for the individual variable \mathbf{x} in the wff \mathbf{A} , φ is any assignment, and ψ is that assignment which agrees with φ off \mathbf{x} , while $\psi \mathbf{x} = \mathcal{V}_\varphi \mathbf{t}$, then $\mathcal{V}_\varphi S_{\mathbf{t}}^{\mathbf{x}} \mathbf{A} = \mathcal{V}_\psi \mathbf{A}$.

Proof: We prove 2302 simultaneously for all proper substitutions θ and assignments φ by induction on the construction of \mathbf{A} .

Case 1. \mathbf{A} is a variable. $\mathcal{V}_\varphi(\theta \mathbf{A}) = (\varphi \circ \theta) \mathbf{A} = \mathcal{V}_{\varphi \circ \theta} \mathbf{A}$ by definition.

Case 2. \mathbf{A} is a constant. $\mathcal{V}_\varphi \theta \mathbf{A} = \mathcal{V}_\varphi \mathbf{A} = \mathcal{V} \mathbf{A} = \mathcal{V}_{\varphi \circ \theta} \mathbf{A}$.

Case 3. \mathbf{A} is $\mathbf{F}^n \mathbf{t}_1 \dots \mathbf{t}_n$, where \mathbf{F}^n is a function or predicate variable or constant, and $\mathbf{t}_1, \dots, \mathbf{t}_n$ are terms.

$$\begin{aligned}\mathcal{V}_\varphi(\theta \mathbf{A}) &= \mathcal{V}_\varphi((\theta \mathbf{F}^n)(\theta \mathbf{t}_1) \dots (\theta \mathbf{t}_n)) = (\mathcal{V}_\varphi \theta \mathbf{F}^n)(\mathcal{V}_\varphi \theta \mathbf{t}_1) \dots (\mathcal{V}_\varphi \theta \mathbf{t}_n) \\ &= (\mathcal{V}_{\varphi \circ \theta} \mathbf{F}^n)(\mathcal{V}_{\varphi \circ \theta} \mathbf{t}_1) \dots (\mathcal{V}_{\varphi \circ \theta} \mathbf{t}_n) \text{ (by inductive hypothesis)} \\ &= \mathcal{V}_{\varphi \circ \theta} \mathbf{F}^n \mathbf{t}_1 \dots \mathbf{t}_n = \mathcal{V}_{\varphi \circ \theta} \mathbf{A}.\end{aligned}$$

Case 4. \mathbf{A} is $\sim \mathbf{B}$.

$$\begin{aligned}\mathcal{V}_\varphi(\theta \sim \mathbf{B}) &= \mathcal{V}_\varphi(\sim \theta \mathbf{B}) = \sim \mathcal{V}_\varphi(\theta \mathbf{B}) \\ &= \sim \mathcal{V}_{\varphi \circ \theta} \mathbf{B} \text{ (by inductive hypothesis)} \\ &= \mathcal{V}_{\varphi \circ \theta} \sim \mathbf{B}\end{aligned}$$

Case 5. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$

$$\begin{aligned}\mathcal{V}_\varphi \theta [\mathbf{B} \vee \mathbf{C}] &= \mathcal{V}_\varphi[(\theta \mathbf{B}) \vee (\theta \mathbf{C})] = \mathcal{V}_\varphi(\theta \mathbf{B}) \vee \mathcal{V}_\varphi(\theta \mathbf{C}) \\ &= \mathcal{V}_{\varphi \circ \theta} \mathbf{B} \vee \mathcal{V}_{\varphi \circ \theta} \mathbf{C} \text{ (by inductive hypothesis)} \\ &= \mathcal{V}_{\varphi \circ \theta} [\mathbf{B} \vee \mathbf{C}]\end{aligned}$$

Case 6. \mathbf{A} is $\forall \mathbf{y} \mathbf{B}$.

Note that if \mathbf{x} is any free variable of \mathbf{B} other than \mathbf{y} , then \mathbf{y} is not free in $\theta \mathbf{x}$, since θ is free on $\forall \mathbf{y} \mathbf{B}$. Let θ' be the substitution which agrees with θ off \mathbf{y} , while $\theta' \mathbf{y} = \mathbf{y}$. Thus $\theta \mathbf{A} = \forall \mathbf{y} \theta' \mathbf{B}$. Clearly θ' is free on \mathbf{B} , so by inductive hypothesis, $\mathcal{V}_{\varphi'} \theta' \mathbf{B} = \mathcal{V}_{\varphi' \circ \theta'} \mathbf{B}$ for any assignment φ' .

Subcase (6a). $\mathcal{V}_\varphi(\theta \mathbf{A}) = \mathbf{F}$.

Thus there is an assignment φ' which agrees with φ off y such that $\mathcal{V}_{\varphi'}\theta'B = F$. (We wish to show that $\mathcal{V}_{\varphi \circ \theta}\forall yB = F$. For this purpose we need an assignment ψ which agrees with $\varphi \circ \theta$ off y such that $\mathcal{V}_\psi B = F$.⁶) Let ψ be that assignment which agrees with $\varphi \circ \theta$ off y , while $\psi y = (\varphi' \circ \theta')y$. If x is any free variable of B other than y , then $\mathcal{V}_{\varphi'}\theta x = \mathcal{V}_{\varphi'}\theta x$ by 2300. Hence $\psi x = (\varphi \circ \theta)x = \mathcal{V}_{\varphi}\theta x = \mathcal{V}_{\varphi'}\theta x = \mathcal{V}_{\varphi'}\theta'x = (\varphi' \circ \theta')x$, so ψ agrees with $(\varphi' \circ \theta')$ on all free variables of B . Thus by 2300, $\mathcal{V}_\psi B = \mathcal{V}_{\varphi' \circ \theta'}B = \mathcal{V}_{\varphi'}\theta'B = F$, so $\mathcal{V}_{\varphi \circ \theta}A = \mathcal{V}_{\varphi \circ \theta}\forall yB = F = \mathcal{V}_{\varphi}\theta A$.

Subcase (6b). $\mathcal{V}_{\varphi \circ \theta}A = F$.

Thus there is an assignment ψ which agrees with $\varphi \circ \theta$ off y such that $\mathcal{V}_\psi B = F$. Let φ' be that assignment which agrees with φ off y , while $\varphi'y = \psi y$. If x is any free variable of B other than y , then $\mathcal{V}_{\varphi}\theta x = \mathcal{V}_{\varphi'}\theta x$ by 2300, so $\psi x = (\varphi \circ \theta)x = \mathcal{V}_{\varphi}\theta x = \mathcal{V}_{\varphi'}(\theta x) = \mathcal{V}_{\varphi'}(\theta'x) = (\varphi' \circ \theta')x$. Also, $\psi y = \varphi'y = \mathcal{V}_{\varphi'}\theta'y = (\varphi' \circ \theta')y$, so ψ agrees with $\varphi' \circ \theta'$ on all free variables of B . Hence by 2300, $\mathcal{V}_{\varphi}\theta'B = \mathcal{V}_{\varphi' \circ \theta'}B = \mathcal{V}_\psi B = F$, so $\mathcal{V}_{\varphi}\theta A = \mathcal{V}_\varphi\forall y\theta'B = F = \mathcal{V}_{\varphi \circ \theta}A$.

If neither subcase (6a) nor (6b) applies, then $\mathcal{V}_{\varphi}\theta A$ and $\mathcal{V}_{\varphi \circ \theta}A$ must both be true, so the proof is complete. ■

2303 Soundness Theorem.

- (a) Every theorem of \mathcal{F} is valid.
- (b) If \mathcal{T} is a first-order theory and \mathcal{M} is a model of \mathcal{T} , then every theorem of \mathcal{T} is valid in \mathcal{M} .
- (c) If \mathcal{G} is a set of wffs and φ simultaneously satisfies \mathcal{G} (in some interpretation) and $\mathcal{G} \vdash B$, then φ satisfies B .

Proof: We first show that if D is any axiom of \mathcal{F} , then $\models D$. This is evident if D is an instance of any of Axiom Schemata 1-3. To deal with Axiom Schemata 4 and 5, let any interpretation, and an arbitrary assignment φ , be given; we shall show that $\mathcal{V}_\varphi D = T$.

Suppose D is an instance of Axiom Schema 4, so D has the form $[\forall x A \supset S_t^x A]$, where t is a term free for x in A . If $\mathcal{V}_\varphi \forall x A = F$, clearly $\mathcal{V}_\varphi D = T$. On the other hand, if $\mathcal{V}_\varphi \forall x A = T$, then $\mathcal{V}_\psi A = T$ for any

⁶It might be natural to try letting $\psi = \varphi' \circ \theta'$. One could then attempt to show that $\varphi' \circ \theta'$ agrees with $\varphi \circ \theta$ off y by arguing that for any variable z distinct from y , $(\varphi' \circ \theta')z = \mathcal{V}_{\varphi'}\theta'z = \mathcal{V}_{\varphi'}\theta z \stackrel{?}{=} \mathcal{V}_\varphi\theta z = (\varphi \circ \theta)z$. However, the step marked $\stackrel{?}{=}$ would not be correct, since y might be free in θz .

assignment ψ which agrees with φ off x . In particular, if $\psi = \varphi \circ S_t^x$, then by the Substitution-Value Theorem $\mathcal{V}_\varphi S_t^x A = \mathcal{V}_\psi A = T$, so $\mathcal{V}_\varphi D = T$.

Suppose D is an instance of Axiom Schema 5, and so has the form $[\forall x [A \vee B] \supset \neg A \vee \forall x B]$, where x is not free in A . If $\mathcal{V}_\varphi \forall x [A \vee B] = F$ or $\mathcal{V}_\varphi A = T$, then $\mathcal{V}_\varphi D = T$. Suppose $\mathcal{V}_\varphi \forall x [A \vee B] = T$ and $\mathcal{V}_\varphi A = F$, and let ψ be any assignment which agrees with φ off x . Then $\mathcal{V}_\psi [A \vee B] = T$, but $\mathcal{V}_\psi A = \mathcal{V}_\varphi A = F$ by 2300, so $\mathcal{V}_\psi B = T$. Hence $\mathcal{V}_\varphi \forall x B = T$, so $\mathcal{V}_\varphi D = T$ in this case also.

It is easy to see that if C and $C \supset D$ are valid [in an interpretation], then D is valid [in that interpretation], and that if C is valid [in an interpretation], then $\forall x C$ is also valid [in that interpretation]. Thus every wff in a proof [in a first-order theory] must be valid [in every model of that theory]. This proves (a) and (b).

To prove (c), if $\mathcal{G} \vdash B$ then there is some finite subset $\{A_1, \dots, A_n\}$ of \mathcal{G} such that $A_1, \dots, A_n \vdash B$. (The case where the subset is empty is trivial, using (a).) Hence by the Deduction Theorem, $\vdash A_1 \supset \dots \supset \neg A_n \supset B$, so by (a) $\mathcal{V}_\varphi [A_1 \supset \dots \supset \neg A_n \supset B] = T$. However, we are given that $\mathcal{V}_\varphi A_1 = \dots = \mathcal{V}_\varphi A_n = T$, so $\mathcal{V}_\varphi B = T$.

REMARK. Note that the wff $[Px \supset \forall x Px]$ is not valid, and hence is not a theorem. Of course, if A is any wff such that $\vdash A$, then $\vdash \forall x A$, so in this case $\vdash A \supset \forall x A$ by Rule P.

2304 Theorem. Every set of sentences which has a model is consistent.

Proof: Suppose S is a set of sentences which has a model M but S is inconsistent. Then if A is any wff, $S \vdash A \wedge \sim A$, so $M \models A \wedge \sim A$ by the Soundness Theorem (2303). This is impossible, so S must be consistent. ■

EXERCISES

X2300. Show that $M \vdash_\varphi \exists x A$ iff there is an assignment ψ which agrees with φ off x such that $M \models_\psi A$.

X2301.

- (a) Find a wff A such that $[A \supset \forall x A]$ is valid.
- (b) Find a wff B such that $[B \supset \forall x B]$ is not valid.
- (c) Prove that if C is any wff and M is an interpretation, then $[C \supset \forall x C]$ is satisfiable in M .

X2302. Is it true that for every wff \mathbf{A} of \mathcal{F} , $\models \mathbf{A}$ iff $\models \mathbf{A} \supset \forall x\mathbf{A}$?

X2303. Are there any valid wffs of the form $\forall x[\mathbf{A} \vee \mathbf{B}] \supset \neg \mathbf{A} \vee \forall x\mathbf{B}$ with x free in \mathbf{A} ?

X2304. Are wffs of the following forms necessarily valid? satisfiable?

- (a) $\exists x[\mathbf{M} \equiv \mathbf{N}] \supset \neg \exists x\mathbf{M} \equiv \exists x\mathbf{N}$
- (b) $\forall x[\mathbf{M} \equiv \mathbf{N}] \equiv \neg \forall x\mathbf{M} \equiv \forall x\mathbf{N}$

X2305. Let $A(x)$ be the wff $[Px \supset \neg Pa \wedge Pb]$. Show that $\models \exists xA(x)$, but there is no term t such that $\models A(t)$.

X2306. Let the wffs A , B , and C be defined as follows:

$$\begin{aligned} A &\text{ is } \exists z[\exists x \forall y Pxyz \supset \exists w Qwz] \\ B &\text{ is } \exists z[\forall y \exists x Pxyz \supset \exists w Qwz] \\ C &\text{ is } \forall z \forall y \exists x Pxyz \supset \exists z \exists w Qwz \end{aligned}$$

For each ordered pair (D, E) of distinct wffs, where D and E are taken from the set $\{A, B, C\}$, decide whether or not $[D \supset E]$ is valid.

X2307. Find a model for the sentence $\forall x[Px \equiv \sim Pf x]$. (Here P and f are constants.)

X2308. For any wff A of \mathcal{F} , let A' be defined as follows: $A' = \mathbf{A}$ if \mathbf{A} is atomic; $(\sim \mathbf{B})' = \sim \mathbf{B}'$; $[\mathbf{B} \vee \mathbf{C}]' = [\mathbf{B}' \vee \mathbf{C}']$; $(\forall x\mathbf{B})' = \exists x\mathbf{B}'$. Is it true that $\vdash \mathbf{A} \supset \mathbf{A}'$ for every wff \mathbf{A} of \mathcal{F} ?

In X2309-X2311, prove or refute the stated conjectures.

X2309. If \mathbf{A} and \mathbf{B} are satisfiable sentences, then the first-order theory with $\{\mathbf{A}, \mathbf{B}\}$ as its set of postulates is consistent.

X2310. If \mathbf{M} , \mathbf{N} , \mathbf{C} , and \mathbf{C}' are wffs of \mathcal{F} such that \mathbf{C}' is obtained from \mathbf{C} by replacing \mathbf{M} by \mathbf{N} at one positive occurrence of \mathbf{M} in \mathbf{C} , and if $\vdash \mathbf{M} \vee \mathbf{N}$, then $\vdash \mathbf{C} \vee \mathbf{C}'$.

X2311. If \mathbf{M} , \mathbf{N} , \mathbf{C} , and \mathbf{C}' are wffs of \mathcal{F} such that \mathbf{C}' is obtained from \mathbf{C} by replacing \mathbf{M} by \mathbf{N} at one positive occurrence of \mathbf{M} in \mathbf{C} , and if $\vdash \mathbf{M} \wedge \mathbf{N}$, then $\vdash \mathbf{C} \wedge \mathbf{C}'$.

Decide which of the wffs below are valid. Prove each valid wff and give a countermodel for each wff which is not valid.

X2312. $\exists y Qywy \equiv \exists z Qzx$

X2313. $[\exists xPx \supset \forall xQx] \supset \forall x[Px \supset Qx]$

X2314. $\forall x[Px \supset Qx] \supset [\exists xPx \supset \forall xQx]$

X2315. $\forall x[Px \supset Qx] \equiv \neg \forall xPx \supset \exists xQx$

X2316. $\forall x[Px \supset Qx] \equiv \neg \forall xPx \supset \forall xQx$

X2317. $\forall x[Qx \supset Rx] \supset \neg \exists x[Px \wedge Qx] \supset \exists x \neg Px \wedge Rx$

X2318. $\exists x \forall y \neg Px \equiv Py$

X2319. $\exists y \forall y Rxy \supset \forall x \exists y Rxy$

X2320. $\exists x \forall y Rxy \equiv \forall y \exists x Rxy$

X2321. $\exists y \forall x[Px \equiv \neg Py \wedge Qx] \supset \exists z \neg Pz \supset Qz$

X2322. $\forall x \forall y \neg Pxy \supset \exists z Pzz$

X2323. $\forall y \neg Pyy \supset \exists z Pzy$

X2324. $\exists x \forall y[Pxy \wedge Qx] \supset \neg \exists y Qy \wedge \exists z Pzz$

§24. Independence

2400 Lemma. If $\vdash A$, then \vee occurs in A .

Proof: It is straightforward to establish by induction on the construction of A that if A is any wff in which \vee does not occur, then A is satisfiable in some interpretation with one individual, and $\sim A$ is satisfiable in some interpretation with one individual. Hence if \vee does not occur in A , by 2301a A is not valid, and so by 2303 is not a theorem. ■

2401 Theorem. The rules of inference and axiom schemata of \mathcal{F} are independent.

Proof: To see that Modus Ponens is independent, note that every theorem of \mathcal{F} which can be proved without using Modus Ponens is obtained from an axiom by zero or more applications of Generalization. Clearly $p \vee \sim p$ (a theorem by 2100) cannot be so obtained, so Modus Ponens is independent.

To see that Generalization is independent, modify the definition of the valuation function V of §23 so that $V_\varphi^M \forall x A$ is always F. With this modified definition, the axioms are valid and Modus Ponens preserves validity, so any theorem provable without using Generalization if valid in the modified

sense. However, $\forall x [p \vee p \supset p]$, which is a theorem of \mathcal{F} by Axiom 1 and Generalization, is not valid in the modified sense, and so cannot be proved without using Generalization. Thus Generalization is an independent rule of inference.

The independence of each of Axiom Schemata 1, 2, and 3 of \mathcal{F} is readily established by a modification of the method used in §13 to establish the independence of the corresponding Axiom Schema of \mathcal{P} . For any wff \mathbf{A} of \mathcal{F} , let \mathbf{A}' be the formula obtained by deleting all quantifiers and terms from \mathbf{A} . (Thus, if \mathbf{A} is the axiom $\sim \forall x P^2 x f^1 y \vee P^2 g^2 y f^1 x f^1 y$, then \mathbf{A}' is $\sim P^2 \vee P^2$.) We may regard \mathbf{A}' as a wff of some system of propositional calculus by including the predicate letters of \mathcal{F} among the propositional variables of that system. To prove the independence of Axiom Schema i (where $i = 1, 2$, or 3) of \mathcal{F} , define a wff \mathbf{A} of \mathcal{F} to have property \mathcal{R}_i iff \mathbf{A}' is special in the sense of §13 when \vee and \sim are interpreted according to Figure 1.9 + i in §13. It is then readily verified that Modus Ponens and Generalization preserve property \mathcal{R}_i , and that each axiom of \mathcal{F} which is not an instance of Axiom Schema i has property \mathcal{R}_i . (Note that if \mathbf{A} is an instance of Axiom Schema 4 or 5, then \mathbf{A}' has the form $\sim \mathbf{B} \vee \mathbf{B}$.) However, as in §13 we can find an instance of Axiom Schema i which does not have property \mathcal{R}_i , so this wff cannot be derived from the other axiom schemata of \mathcal{F} . Hence Axiom Schema i is independent for $i = 1, 2$, and 3 .

To prove the independence of Axiom Schema 4, for each wff \mathbf{A} , let \mathbf{A}' be the result of replacing $\forall x$ by $\exists x$ (i.e., $\sim \forall x \sim$) everywhere in \mathbf{A} . Say that \mathbf{A} has property \mathcal{R}_4 iff $\models \mathbf{A}'$. It is easy to see that all instances of Axiom Schemata 1, 2, 3, and 5 have property \mathcal{R}_4 , and that Modus Ponens and Generalization preserve property \mathcal{R}_4 . However, the instance $[\forall x P^1 x \supset P^1 x]$ of Axiom Schema 4 does not have property \mathcal{R}_4 , since $\mathcal{V}_\varphi^\mathcal{M} [\exists x P^1 x \supset P^1 x] = F$ when \mathcal{M} is an interpretation with the integers as domain of individuals, φP^1 is the property of being positive, and $\varphi x = -2$. Thus Axiom Schema 4 is independent. (This argument suggests that one of the reasons for including Axiom Schema 4 among the axioms of \mathcal{F} is to prevent $\forall x$ from being soundly interpreted as meaning “there exists x such that”.)

To prove the independence of Axiom Schema 5, let \mathbf{A}' be the wff obtained from \mathbf{A} by replacing each wf part of \mathbf{A} of the form $\forall x \mathbf{B}$, where \mathbf{B} contains no occurrences of \vee , by the wff $[p \wedge \sim p]$. (This may be done sequentially, treating the quantifiers of \mathbf{A} in the left-to-right order of their occurrences in \mathbf{A} .) Say that \mathbf{A} has property \mathcal{R}_5 iff $\vdash \mathbf{A}$ and $\models \mathbf{A}'$. Clearly Generalization preserves property \mathcal{R}_5 , since if $\vdash \mathbf{A}$ and $\models \mathbf{A}'$, then $\vdash \forall x \mathbf{A}$ and $\models \forall x (\mathbf{A}')$ (by 2301c), but \mathbf{A} contains \vee by 2400, so $(\forall x \mathbf{A})'$ is $\forall x (\mathbf{A}')$. Modus Ponens also preserves property \mathcal{R}_5 , since $[\mathbf{A} \supset \mathbf{B}]'$ is $[\mathbf{A}' \supset \mathbf{B}']$.

All instances of Axiom Schemata 1-3 clearly have property \mathcal{R}_5 . It can be seen that $[\forall x \mathbf{A} \supset S_t^x \mathbf{A}]'$ is $[[p \wedge \sim p] \supset S_t^x (\mathbf{A}')]$ if \vee does not occur in \mathbf{A} , and otherwise $\forall x (\mathbf{A}') \supset S_t^x (\mathbf{A}')$; moreover, if t is free for x in \mathbf{A} , it is also free for x in \mathbf{A}' . Hence all instances of Axiom Schema 4 have property \mathcal{R}_5 . However, if \mathbf{A} is the instance $[\forall x [r \vee Q^1 x] \supset \neg r \vee \forall x Q^1 x]$ of Axiom Schema 5, then $\mathcal{V}_\varphi \mathbf{A}' = \mathcal{V}_\varphi [\forall x [r \vee Q_1 x] \supset \neg r \vee \neg p \wedge \sim p] = F$ whenever $\varphi r = F$ and φQ^1 is the predicate which is true of all individuals (in a given interpretation). Thus, \mathbf{A} does not have property \mathcal{R}_5 , so Axiom Schema 5 is independent. ■

We can now show, as promised in the remark following metatheorem 2113, that in any proof of $\forall y Py \supset \forall x Px$, there must be an application of the rule of Generalization in which x is the variable generalized upon. For this purpose, modify the definition of \mathcal{V}_φ in §23 so that for any wff \mathbf{A} , $\mathcal{V}_\varphi \forall x \mathbf{A} = F$. However, if y is any individual variable other than x , $\mathcal{V}_\varphi \forall y \mathbf{A}$ is defined as before. With this modified definition, it can easily be checked that the axioms remain valid, Modus Ponens preserves validity, and Generalization with respect to any variable other than x preserves validity. However, $\forall y Py \supset \forall x Px$ is not valid under this definition, as can be seen by assigning P as value the set of all individuals in some interpretation, so that $\mathcal{V}_\varphi \forall y Py = T$. Hence any proof of $\forall y Py \supset \forall x Px$ in \mathcal{F} must contain a generalization upon x .

EXERCISES

X2400. Show that if $\vdash \mathbf{A}$, then \sim occurs in \mathbf{A} .

X2401. Give a careful inductive definition of \mathbf{A}' in the proof of the independence of Axiom Schema 5. Then give a careful inductive proof that $(S_t^x \mathbf{A})' = S_t^x (\mathbf{A}')$ for any term t .

§25. Abstract Consistency and Completeness

We start this section with a number of incidental definitions and results which will be useful later.

DEFINITION. A *literal* is a wff which is atomic, or is the negation of an atomic wff.

2500 Proposition. Every wff of \mathcal{F} is a literal or is of one of the forms $[A \vee B]$, $\forall x A$, $\sim\sim A$, $\sim[A \vee B]$, $\sim\forall x A$.

Proof: Clearly every wff which is not atomic must have one of the forms $\sim C$, $[A \vee B]$, or $\forall x A$. If it has the form $\sim C$, then C must also be atomic or have one of these forms. ■

Recall that a set \mathcal{G} of wffs is consistent iff there is a wff A such that not $\mathcal{G} \vdash A$.

DEFINITION. The set \mathcal{H} of wffs is *consistent with \mathcal{G}* iff $\mathcal{G} \cup \mathcal{H}$ is consistent. We say that A is *consistent with \mathcal{G}* iff $\mathcal{G} \cup \{A\}$ is consistent.

2501 Lemma. Let \mathcal{G} be any set of wffs. $\{A_1, \dots, A_n\}$ is inconsistent with \mathcal{G} iff $\mathcal{G} \vdash \sim A_1 \vee \dots \vee \sim A_n$. The wff A is inconsistent with \mathcal{G} iff $\mathcal{G} \vdash \sim A$.

Proof: If $\{A_1, \dots, A_n\}$ is inconsistent with \mathcal{G} , then $\mathcal{G}, A_1, \dots, A_n \vdash \sim A_1$ so $\mathcal{G} \vdash A_1 \supset \dots \supset A_n \supset \sim A_1$ by the Deduction Theorem. Hence $\mathcal{G} \vdash \sim A_1 \vee \dots \vee \sim A_n$ by Rule P.

If $\mathcal{G} \vdash \sim A_1 \vee \dots \vee \sim A_n$ then $\mathcal{G}, A_1, \dots, A_n \vdash \sim A_1 \vee \dots \vee \sim A_n$ by 2113. Also, $\mathcal{G}, A_1, \dots, A_n \vdash A_i$ for $1 \leq i \leq n$ so if B is any wff, $\mathcal{G}, A_1, \dots, A_n \vdash B$ by Rule P. Hence $\{A_1, \dots, A_n\}$ is inconsistent with \mathcal{G} . ■

Note that the proof above applies when $n = 1$.

To simplify terminology, until further notice we shall restrict our attention to formulations of \mathcal{F} in which there are no variables except individual variables. Thus all predicate and function symbols will be constants, and all closed wffs will be sentences. Various formulations of \mathcal{F} will differ from each other only by having different sets of constants. It will be obvious how to transfer results about these formulations of \mathcal{F} to formulations with other types of variables.

DEFINITIONS. We shall refer to the set of wffs of a logistic system \mathcal{F} as the *language of \mathcal{F}* , and denote this set by $\mathcal{L}(\mathcal{F})$. If \mathcal{F}^1 and \mathcal{F}^2 are logistic systems, we shall say that \mathcal{F}^2 is an *expansion* of \mathcal{F}^1 iff $\mathcal{L}(\mathcal{F}^1) \subseteq \mathcal{L}(\mathcal{F}^2)$, and a *proper expansion* iff $\mathcal{L}(\mathcal{F}^1) \subset \mathcal{L}(\mathcal{F}^2)$. For the present we are concerned only with formulations of the first-order system \mathcal{F} , so \mathcal{F}^2 is an expansion of \mathcal{F}^1 iff every constant of \mathcal{F}^1 is a constant of \mathcal{F}^2 . Naturally, when one expands a formulation of \mathcal{F} by adding new constants, one also

adds new instances of the logical axiom schemata to the list of axioms.

If \mathcal{F}^1 and \mathcal{F}^2 are first-order theories, we shall say that \mathcal{F}^2 is an *extension* of \mathcal{F}^1 iff $\mathcal{L}(\mathcal{F}^1) \subseteq \mathcal{L}(\mathcal{F}^2)$ and every theorem of \mathcal{F}^1 is a theorem of \mathcal{F}^2 . The latter condition is satisfied iff every postulate of \mathcal{F}^1 is a theorem of \mathcal{F}^2 . Of course, postulates of \mathcal{F}^1 need not be postulates of \mathcal{F}^2 .

Let \mathcal{F}^1 and \mathcal{F}^2 be first order theories. We say that \mathcal{F}^2 is a *conservative extension* of \mathcal{F}^1 iff for every wff C of $\mathcal{L}(\mathcal{F}^1)$, $\vdash_{\mathcal{F}^1} C$ iff $\vdash_{\mathcal{F}^2} C$. Clearly a conservative extension is an extension.

Logistic systems \mathcal{F}^1 and \mathcal{F}^2 are *equivalent* iff each is an extension of the other. In this case, they have the same wffs and the same theorems, and each is a conservative extension of the other.

An example of a conservative extension might be obtained by letting \mathcal{F}^1 be a formalization of the mathematics used in a certain scientific theory, and creating \mathcal{F}^2 by adding to \mathcal{F}^1 constants to serve as names various scientific concepts, and adding axioms to express the definitions and laws of the scientific theory. One would expect that if a theorem of \mathcal{F}^2 had purely mathematical content (i.e., was a wff of \mathcal{F}^1), its proof would not require the assumptions about the physical world expressed by the laws of the scientific theory, or the definitions of the concepts of that theory, so it should be provable in \mathcal{F}^1 . That is, \mathcal{F}^2 should be a conservative extension of \mathcal{F}^1 .

2502 Proposition. Let \mathcal{F}^1 and \mathcal{F}^2 be first order theories, with \mathcal{F}^2 a conservative extension of \mathcal{F}^1 . Then \mathcal{F}^2 is consistent iff \mathcal{F}^1 is consistent.

Proof: Left to the reader.

If \mathcal{G} is a set of sentences of \mathcal{F} , we shall denote by $\mathcal{F} \cup \mathcal{G}$ the first order theory with the set \mathcal{G} of postulates. (Thus the set of axioms of this theory is the union of \mathcal{G} with the set of axioms \mathcal{F} .)

2503 Proposition. Let \mathcal{F}^1 and \mathcal{F}^2 be formulations of \mathcal{F} , so that \mathcal{F}^2 is an expansion of \mathcal{F}^1 obtained by adding additional individual constants to the set of constants of \mathcal{F}^1 . Let \mathcal{G} be any set of sentences of \mathcal{F}^1 . Then

- (1) $\mathcal{F}^2 \cup \mathcal{G}$ is a conservative extension of $\mathcal{F}^1 \cup \mathcal{G}$.
- (2) \mathcal{G} is consistent in \mathcal{F}^1 iff it is consistent in \mathcal{F}^2 .

Proof: Clearly $\mathcal{F}^2 \cup \mathcal{G}$ is an extension of $\mathcal{F}^1 \cup \mathcal{G}$. If \mathcal{P} is a proof in $\mathcal{F}^2 \cup \mathcal{G}$ of a wff C of \mathcal{F}^1 , then one can replace the individual constants of $\mathcal{L}(\mathcal{F}^2) - \mathcal{L}(\mathcal{F}^1)$ in \mathcal{P} by new individual variables to obtain a proof of C in $\mathcal{F}^1 \cup \mathcal{G}$.

Clearly \mathcal{G} is consistent in \mathcal{F} iff $\mathcal{F} \cup \mathcal{G}$ is consistent, so (2) follows from (1) and 2502. ■

We say that a set is *countable* iff it is finite or denumerably infinite. Since zero is finite, a countable set may be empty.

Customarily we use formulations of \mathcal{F} with countably many constants. Indeed, one never actually writes down more than finitely many wffs, so a countably infinite set of constants provides an inexhaustible supply of new constants for those who wish to use first-order logic as a vehicle for formalizing theories and proving theorems. Nevertheless, for certain model-theoretic purposes it is useful to consider, as objects of mathematical study, formulations of first-order logic with uncountably many constants. Therefore we shall impose no restrictions on the cardinality of the set of constants. We always assume that there are denumerably many individual variables.

We denote the cardinality of any set or ordinal number S by $\text{card}(S)$. (The cardinality of an ordinal is the cardinality of the set of ordinals which precede it.) Since every formula is a finite sequence of primitive symbols, it is easy to see that $\text{card}(\mathcal{L}(\mathcal{F}))$ is the cardinality of the set of primitive symbols of \mathcal{F} . Formulations of \mathcal{F} with countably many constants have \aleph_0 wffs, and are called countable formulations. (Indeed, one can enumerate the primitive symbols of a countable formulation of \mathcal{F} , so that each primitive symbol K is assigned a positive integer ' K '. Then to each formula $K_0 \dots K_n$ one can assign the number $\prod_{i=0}^n P_i^{K_i}$, where P_i is the i th odd prime (as defined for Theorem 6522 in §65). Then each formula is assigned a number, and distinct formulas have distinct numbers.)

The reader who does not wish to be bothered with cardinality considerations may focus entirely on countable formulations of \mathcal{F} . Indeed, this is advisable on first encounter with the material in this section. We phrase the proofs so as to apply to formulations of \mathcal{F} of arbitrary cardinality because very little extra effort is required to achieve this generality.

An interpretation $\langle \mathcal{D}, \mathcal{J} \rangle$ for \mathcal{F} is said to be *frugal* iff $\text{card}(\mathcal{D}) \leq \text{card}(\mathcal{L}(\mathcal{F}))$. If \mathcal{F} is countable, this means that the domain of individuals is countable, and a frugal interpretation [model] is called a countable interpretation [model].

We shall show in Theorem 2508 that every consistent set of sentences has a frugal model. Note that if \mathcal{F} has uncountably many constants, a consistent set of sentences of \mathcal{F} need not have a *countable* model. For example, consider the set \mathcal{S} which contains (1) the sentence $\forall x E xx$ and (2) all sentences of the form $\sim Ecd$, where c and d are distinct individual constants. In any model of \mathcal{S} , the predicate constant E denotes a reflexive relation, such as

equality. Therefore, it is easy to see that if \mathcal{M} is any model for \mathcal{S} , and c and d are distinct constants, then $\mathcal{V}^{\mathcal{M}}c \neq \mathcal{V}^{\mathcal{M}}d$. For if $\mathcal{V}^{\mathcal{M}}c = \mathcal{V}^{\mathcal{M}}d$, and we let $\psi x = \mathcal{V}^{\mathcal{M}}c$, then by (1) $T = \mathcal{V}_{\psi}^{\mathcal{M}}Exx = (\mathcal{V}^{\mathcal{M}}E)(\psi x)(\psi x) = (\mathcal{V}^{\mathcal{M}}E)(\mathcal{V}^{\mathcal{M}}c)(\mathcal{V}^{\mathcal{M}}d) = \mathcal{V}^{\mathcal{M}}Ecd$, contradicting (2). Thus, any model of \mathcal{S} must have at least as many individuals as there are individual constants in \mathcal{F} .

Our main objective in this section is to prove Theorem 2508 and its consequences, including Gödel's Completeness Theorem (2510). Clearly, the concept of consistency is the fundamental syntactic notion in Theorem 2508. It was noticed by Smullyan [Smullyan, 1963] that the proof of Theorem 2508 uses only certain properties of the notion of consistency, and that by treating consistency in an abstract way, with very little extra effort one can obtain a generalization (2507 Smullyan's Unifying Principle) of Theorem 2508 which yields a variety of important metatheorems. The power of Smullyan's Unifying Principle is a consequence of the fact that various logical questions can be formulated as questions about consistency relative to special logical systems. Our only use of Smullyan's Unifying Principle in this chapter will be in the proof of Theorem of 2508, but other examples of its usefulness will be found in later chapters. In §25A we give a direct proof of Theorem 2508 for countable formulations of \mathcal{F} . The reader who is not interested in the more general treatment below may now proceed to read §25A, and then go directly to Theorem 2509.

We shall formulate an abstract notion of consistency by considering various properties possessed by the class⁷ of consistent sets of sentences. Two such properties are defined below:

DEFINITION. Let Γ be a class of sets.

- (1) Γ is *closed under subsets* iff for all sets S_1 and S_2 , if $S_1 \subseteq S_2$ and $S_2 \in \Gamma$, then $S_1 \in \Gamma$.
- (2) Γ is of *finite character* iff for every set S , $S \in \Gamma$ iff every finite subset of S is a member of Γ .

2504 Lemma. If Γ is a class of sets which is of finite character, then Γ is closed under subsets. ■

Proof: Suppose $S_1 \subseteq S_2$ and $S_2 \in \Gamma$. Every finite subset of S_1 is a subset of S_2 , and so is in Γ , so $S_1 \in \Gamma$. ■

⁷We use the word *class* as a synonym for *collection* or *set*, without any of the connotations associated with the word in axiomatic set theory.

A *variable-free* term is a term in which no variables occur.

DEFINITION. We shall say that Γ is an *abstract consistency class* iff Γ is a class of sets of sentences, Γ is closed under subsets, and for all sets \mathcal{S} in Γ , the following conditions hold for all wffs \mathbf{A} and \mathbf{B} :

- (a) If \mathbf{A} is atomic, then $\mathbf{A} \notin \mathcal{S}$ or $\sim \mathbf{A} \notin \mathcal{S}$.
- (b) If $\sim \sim \mathbf{A} \in \mathcal{S}$, then $\mathcal{S} \cup \{\mathbf{A}\} \in \Gamma$.
- (c) If $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{S}$, then $\mathcal{S} \cup \{\mathbf{A}\} \in \Gamma$ or $\mathcal{S} \cup \{\mathbf{B}\} \in \Gamma$.
- (d) If $\sim [\mathbf{A} \vee \mathbf{B}] \in \mathcal{S}$, then $\mathcal{S} \cup \{\sim \mathbf{A}, \sim \mathbf{B}\} \in \Gamma$.
- (e) If $\forall x \mathbf{A} \in \mathcal{S}$, then for each variable-free term t , $\mathcal{S} \cup \{S_t^x \mathbf{A}\} \in \Gamma$.
- (f) If $\sim \forall x \mathbf{A} \in \mathcal{S}$, and c is any individual constant which does not occur in any sentence of \mathcal{S} , then $\mathcal{S} \cup \{\sim S_c^x \mathbf{A}\} \in \Gamma$.

It is easy to see that the class of sets of sentences which have models is an abstract consistency class. Another example, which justifies the terminology, is provided by the following proposition.

2505 Proposition. $\{\mathcal{S} \mid \mathcal{S}$ is a consistent set of sentences of $\mathcal{F}\}$ is an abstract consistency class.

Proof: The class is closed under subsets, since if a set \mathcal{S} is consistent, then every subset of \mathcal{S} is consistent. Clearly if \mathcal{S} is consistent, no wff and its negation can be in \mathcal{S} . This establishes (a).

Note that if $\vdash C \supset D_1 \wedge \dots \wedge D_n$ (where $n \geq 1$) and $C \in \mathcal{S}$ and \mathcal{S} is consistent, then $\mathcal{S} \cup \{D_1, \dots, D_n\}$ must be consistent. For if $\mathcal{S} \cup \{D_1, \dots, D_n\}$ is inconsistent, then $\mathcal{S} \vdash \sim D_1 \vee \dots \vee \sim D_n$ by Lemma 2501, so $\mathcal{S} \vdash \sim C$ by Rule P, so \mathcal{S} is inconsistent. This establishes (b), (d), and (e).

If \mathcal{S} is consistent but $\mathcal{S} \cup \{\mathbf{A}\}$ and $\mathcal{S} \cup \{\mathbf{B}\}$ are both inconsistent, then $\mathcal{S} \vdash \sim \mathbf{A}$ and $\mathcal{S} \vdash \sim \mathbf{B}$ by 2501, so $\mathcal{S} \vdash \sim [\mathbf{A} \vee \mathbf{B}]$ by Rule P, so $[\mathbf{A} \vee \mathbf{B}] \notin \mathcal{S}$. This establishes (c).

To establish (f), suppose \mathcal{S} is consistent but $\mathcal{S} \cup \{\sim S_c^x \mathbf{A}\}$ is inconsistent, where c is an individual constant which does not occur in \mathcal{S} or \mathbf{A} . Then $\mathcal{S} \vdash S_c^x \mathbf{A}$ by 2501 and Rule P, so there is a finite subset \mathcal{H} of \mathcal{S} such that $\mathcal{H} \vdash S_c^x \mathbf{A}$. Given a proof \mathcal{P} of $S_c^x \mathbf{A}$ from \mathcal{H} , let y be an individual variable which does not occur in \mathcal{P} or in \mathcal{H} . Since c does not occur in \mathcal{H} or \mathbf{A} , it can be seen that $S_y^c \mathcal{P}$ is a proof of $S_y^x \mathbf{A}$ from \mathcal{H} , so $\mathcal{H} \vdash S_y^x \mathbf{A}$. Hence $\mathcal{H} \vdash \forall y S_y^x \mathbf{A}$ by Generalization, and $\mathcal{H} \vdash \forall x \mathbf{A}$ by the α Rule, so $\mathcal{S} \vdash \sim \sim \forall x \mathbf{A}$, so $\sim \forall x \mathbf{A} \notin \mathcal{S}$. ■

2506 Proposition. Let Γ be an abstract consistency class. Let $\Delta = \{\mathcal{S} \mid \text{every finite subset of } \mathcal{S} \text{ is in } \Gamma\}$. Then $\Gamma \subseteq \Delta$ and Δ is an abstract consistency class of finite character.

*Proof.*⁸ To see that $\Gamma \subseteq \Delta$, suppose $\mathcal{S} \in \Gamma$. Then every finite subset of \mathcal{S} is in Γ since Γ is closed under subsets, so $\mathcal{S} \in \Delta$.

Next, let us show that Δ is of finite character. Suppose $\mathcal{S} \in \Delta$ and \mathcal{H} is any finite subset of \mathcal{S} . Then by the definition of Δ , all finite subsets of \mathcal{H} are in Γ , so $\mathcal{H} \in \Delta$. Thus all finite subsets of \mathcal{S} are in Δ whenever \mathcal{S} is in Δ . On the other hand, suppose all finite subsets of \mathcal{S} are in Δ . Then by the definition of Δ the finite subsets of \mathcal{S} are also in Γ , so $\mathcal{S} \in \Delta$. Thus Δ is of finite character.

Now we show that Δ is an abstract consistency class. By Lemma 2504 it is closed under subsets. Let $\mathcal{S} \in \Delta$.

(a) Suppose there is an atom \mathbf{A} such that $\{\mathbf{A}, \sim \mathbf{A}\} \subseteq \mathcal{S}$. Then $\{\mathbf{A}, \sim \mathbf{A}\} \in \Gamma$, contradicting (a) for Γ .

(b), (d), (e), (f) We handle these conditions simultaneously, since each has the form “If $\mathbf{C} \in \mathcal{S}$, then $\mathcal{S} \cup \mathcal{C} \in \Delta$ ”, where \mathbf{C} is a wff and \mathcal{C} is an appropriate finite set of wffs; for example in case (e) \mathbf{C} is $\forall x \mathbf{A}$, and \mathcal{C} is $\{S_t^x \mathbf{A}\}$ for some variable-free term t . So suppose $\mathbf{C} \in \mathcal{S}$, and let \mathcal{H} be any finite subset of $\mathcal{S} \cup \mathcal{C}$. Let $\mathcal{S}_1 = (\mathcal{H} - \mathcal{C}) \cup \{\mathbf{C}\}$. \mathcal{S}_1 is a finite subset of \mathcal{S} , so $\mathcal{S}_1 \in \Gamma$. Since Γ is an abstract consistency class and $\mathbf{C} \in \mathcal{S}_1$, we have $\mathcal{S}_1 \cup \mathcal{C} \in \Gamma$. $\mathcal{H} \subseteq \mathcal{S}_1 \cup \mathcal{C}$ and Γ is closed under subsets, so $\mathcal{H} \in \Gamma$. Thus every finite subset \mathcal{H} of $\mathcal{S} \cup \mathcal{C}$ is in Γ , so $\mathcal{S} \cup \mathcal{C} \in \Delta$. (Of course, when dealing with condition (f), we can use the fact that if an individual constant does not occur in any sentence of \mathcal{S} , then it does not occur in any sentence of \mathcal{S}_1 .)

(c) Suppose $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{S}$ but $\mathcal{S} \cup \{\mathbf{A}\} \notin \Delta$ and $\mathcal{S} \cup \{\mathbf{B}\} \notin \Delta$. Then there are finite subsets \mathcal{S}_1 and \mathcal{S}_2 of \mathcal{S} such that $\mathcal{S}_1 \cup \{\mathbf{A}\} \notin \Gamma$ and $\mathcal{S}_2 \cup \{\mathbf{B}\} \notin \Gamma$ (since all finite subsets of \mathcal{S} are in Γ). Let $\mathcal{S}_3 = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \{[\mathbf{A} \vee \mathbf{B}]\}$. \mathcal{S}_3 is a finite subset of \mathcal{S} , so $\mathcal{S}_3 \in \Gamma$. Since Γ is an abstract consistency class and $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{S}_3$, we have $\mathcal{S}_3 \cup \{\mathbf{A}\} \in \Gamma$ or $\mathcal{S}_3 \cup \{\mathbf{B}\} \in \Gamma$. Γ is closed under subsets, so $\mathcal{S}_1 \cup \{\mathbf{A}\} \in \Gamma$ or $\mathcal{S}_2 \cup \{\mathbf{B}\} \in \Gamma$. This is a contradiction, so we conclude that if $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{S}$, then $\mathcal{S} \cup \{\mathbf{A}\} \in \Delta$ or $\mathcal{S} \cup \{\mathbf{B}\} \in \Delta$. ■

⁸Note that if Γ is of finite character then $\Delta = \Gamma$, and this proposition is vacuous. Since many abstract consistency classes can easily be seen to be of finite character, some readers may wish to postpone reading the proof of this proposition.

DEFINITION. Let \mathcal{S} be a set of wffs of \mathcal{F} . \mathcal{S} is *pure* iff no wff of \mathcal{S} contains an individual constant. \mathcal{S} is *sufficiently pure* in \mathcal{F} iff there is a set of cardinality $\text{card}(\mathcal{L}(\mathcal{F}))$ of individual constants of \mathcal{F} which do not occur in any wff of \mathcal{S} .

2507 Smullyan's Unifying Principle [Smullyan, 1963], [Smullyan, 1968, Chapter VI]. If Γ is an abstract consistency class and \mathcal{S} is a sufficiently pure set of sentences and $\mathcal{S} \in \Gamma$, then \mathcal{S} has a frugal model.

Proof: The theorem is trivial if \mathcal{S} is empty, so we may assume \mathcal{S} is non-empty. As in Proposition 2506, let Δ be an abstract consistency class of finite character such that $\Gamma \subseteq \Delta$. (To make the argument seem less abstract, it is suggested that the reader pretend that " $\mathcal{H} \in \Delta$ " means " \mathcal{H} is consistent" when reading the proof for the first time.) Let $K = \text{card}(\mathcal{L}(\mathcal{F}))$, and let κ be the initial ordinal of cardinality K . Well-order the sentences of \mathcal{F} , and for each ordinal $\tau < \kappa$, let \mathbf{S}_τ be the τ -th sentence in this well-ordering.

Let \mathcal{C} be a well-ordered set of cardinality K of individual constants which do not occur in sentences of \mathcal{S} . By transfinite induction, for each ordinal $\tau \leq \kappa$ we define a set \mathcal{S}_τ of sentences so that $\mathcal{S}_\sigma \subseteq \mathcal{S}_\tau$ whenever $\sigma \leq \tau$, and for each τ there is a finite cardinal n such that at most $n + \text{card}(\tau)$ constants from \mathcal{C} occur in the sentences of \mathcal{S}_τ . $\mathcal{S}_0 = \mathcal{S}$. If τ is a limit ordinal, $\mathcal{S}_\tau = \bigcup_{\sigma < \tau} \mathcal{S}_\sigma$. The definition for successor ordinals $\tau + 1$ is by cases:

- (1) If $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\} \notin \Delta$, then $\mathcal{S}_{\tau+1} = \mathcal{S}_\tau$.
- (2) If $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\} \in \Delta$, and \mathbf{S}_τ is not of the form $\sim \forall x \mathbf{A}$, then $\mathcal{S}_{\tau+1} = \mathcal{S}_\tau \cup \{\mathbf{S}_\tau\}$.
- (3) If $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\} \in \Delta$ and \mathbf{S}_τ is $\sim \forall x \mathbf{A}$, then $\mathcal{S}_{\tau+1} = \mathcal{S}_\tau \cup \{\sim \forall x \mathbf{A}, \sim S_c^x \mathbf{A}\}$, where c is the first individual constant from \mathcal{C} which does not occur in any sentence of $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\}$. (Note that there must be such a constant.)

It is readily seen by induction that the condition on the number of constants from \mathcal{C} in \mathcal{S}_τ is satisfied.

We next show by induction that $\mathcal{S}_\tau \in \Delta$ for each $\tau \leq \kappa$. $\mathcal{S}_0 = \mathcal{S} \in \Gamma \subseteq \Delta$. If τ is a limit ordinal, let \mathcal{H} be any finite subset of \mathcal{S}_τ . Then there is a $\sigma < \tau$ such that $\mathcal{H} \subseteq \mathcal{S}_\sigma$, and $\mathcal{S}_\sigma \in \Delta$ by inductive hypothesis, so $\mathcal{H} \in \Delta$ since Δ is closed under subsets. Thus every finite subset of \mathcal{S}_τ is in Δ , so $\mathcal{S}_\tau \in \Delta$. Also, if $\mathcal{S}_\tau \in \Delta$, then $\mathcal{S}_{\tau+1} \in \Delta$ in all cases.

Let $\mathcal{U} = \mathcal{S}_\kappa$. Clearly \mathcal{U} is a set of sentences and $\mathcal{S} \subseteq \mathcal{U}$. We assert that \mathcal{U} has the following properties for all wffs \mathbf{A} and \mathbf{B} :

- (a) There is no atomic sentence \mathbf{A} such that $\mathbf{A} \in \mathcal{U}$ and $\sim \mathbf{A} \in \mathcal{U}$.
- (b) If $\sim \sim \mathbf{A} \in \mathcal{U}$, then $\mathbf{A} \in \mathcal{U}$.
- (c) If $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{U}$, then $\mathbf{A} \in \mathcal{U}$ or $\mathbf{B} \in \mathcal{U}$.
- (d) If $\sim [\mathbf{A} \vee \mathbf{B}] \in \mathcal{U}$, then $\sim \mathbf{A} \in \mathcal{U}$ and $\sim \mathbf{B} \in \mathcal{U}$.
- (e) If $\forall x \mathbf{A} \in \mathcal{U}$, then $S_t^x \mathbf{A} \in \mathcal{U}$ for each variable-free term t .
- (f) If $\sim \forall x \mathbf{A} \in \mathcal{U}$, then there is an individual constant c such that
 $\sim S_c^x \mathbf{A} \in \mathcal{U}$.

(a) is true since $\mathcal{U} \in \Delta$. Next note that if \mathbf{D} is a sentence such that $\mathcal{U} \cup \{\mathbf{D}\} \in \Delta$, then $\mathbf{D} \in \mathcal{U}$. For there is an ordinal τ such that $S_\tau = \mathbf{D}$, and $S_\tau \cup \{\mathbf{D}\} \subseteq \mathcal{U} \cup \{\mathbf{D}\} \in \Delta$, so $S_\tau \cup \{\mathbf{D}\} \in \Delta$ since Δ is closed under subsets. Hence $\mathbf{D} \in S_{\tau+1} \subseteq \mathcal{U}$, so $\mathbf{D} \in \mathcal{U}$. Now to prove (b), suppose $\sim \sim \mathbf{A} \in \mathcal{U}$. Then $\mathcal{U} \cup \{\mathbf{A}\} \in \Delta$ (since Δ is an abstract consistency class), so $\mathbf{A} \in \mathcal{U}$. Similar arguments establish (d) and (e). To prove (c), suppose $[\mathbf{A} \vee \mathbf{B}] \in \mathcal{U}$. Then $\mathcal{U} \cup \{\mathbf{A}\} \in \Delta$ or $\mathcal{U} \cup \{\mathbf{B}\} \in \Delta$, so $\mathbf{A} \in \mathcal{U}$ or $\mathbf{B} \in \mathcal{U}$. To prove (f), suppose $\sim \forall x \mathbf{A} \in \mathcal{U}$. Choose τ such that $S_\tau = \sim \forall x \mathbf{A}$. Then $S_\tau \cup \{\sim \forall x \mathbf{A}\} \subseteq \mathcal{U} \in \Delta$, so $S_\tau \cup \{\sim \forall x \mathbf{A}\} \in \Delta$. Thus $S_{\tau+1} = S_\tau \cup \{\sim \forall x \mathbf{A}, \sim S_c^x \mathbf{A}\}$ for some individual constant c , so $\sim S_c^x \mathbf{A} \in S_{\tau+1} \subseteq \mathcal{U}$.

We next define an interpretation $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$. Let \mathcal{D} be the set of variable-free terms of \mathcal{F} . Thus, $\mathcal{C} \subseteq \mathcal{D}$, so $\text{card}(\mathcal{D}) = \text{card}(\mathcal{L}(\mathcal{F}))$, and the interpretation will be frugal. For each function constant f^n , if t_1, \dots, t_n are in \mathcal{D} , let $(\mathcal{J}f^n)t_1, \dots, t_n = f^n t_1 \dots t_n$, and if a is an individual constant, let $\mathcal{J}a = a$. (Thus variable-free terms denote themselves.) For each predicate constant P^n of \mathcal{F} , if t_1, \dots, t_n are in \mathcal{D} , let $(\mathcal{J}P^n)t_1, \dots, t_n = T$ iff $P^n t_1 \dots t_n \in \mathcal{U}$. Similarly, if p is a propositional constant, $\mathcal{J}p = T$ iff $p \in \mathcal{U}$.

Clearly $\mathcal{V}^\mathcal{M}t = t$ for each variable-free term t , and for each atomic sentence E , $\mathcal{M} \models E$ iff $E \in \mathcal{U}$. We show that by induction on the number of occurrences of \sim , \vee , and \forall in E that if $E \in \mathcal{U}$, then $\mathcal{M} \models E$. (Note Exercise X2500). We consider the following cases:

- (a) E is $\sim A$ for some atomic sentence A . Since $E \in \mathcal{U}$, A cannot be in \mathcal{U} , so $\mathcal{V}^\mathcal{M}A = F$. Hence $\mathcal{M} \models E$.
- (b) E is $\sim \sim A$. Then $A \in \mathcal{U}$, so $\mathcal{M} \models A$ by inductive hypothesis, so $\mathcal{M} \models E$.
- (c) E is $[A \vee B]$. Then $A \in \mathcal{U}$ or $B \in \mathcal{U}$, so by inductive hypothesis $\mathcal{M} \models A$ or $\mathcal{M} \models B$. Hence $\mathcal{M} \models E$ in either case.

- (d) \mathbf{E} is $\sim [\mathbf{A} \vee \mathbf{B}]$. Then $\sim \mathbf{A} \in \mathcal{U}$ and $\sim \mathbf{B} \in \mathcal{U}$ so $\mathcal{M} \models \sim \mathbf{A}$ and $\mathcal{M} \models \sim \mathbf{B}$ by inductive hypothesis, so $\mathcal{V}^{\mathcal{M}}\mathbf{A} = \mathbf{F} = \mathcal{V}^{\mathcal{M}}\mathbf{B}$, so $\mathcal{V}^{\mathcal{M}}[\mathbf{A} \vee \mathbf{B}] = \mathbf{F}$ and $\mathcal{M} \models \mathbf{E}$.
- (e) \mathbf{E} is $\forall \mathbf{x}\mathbf{A}$. Let ψ be any assignment, and let $\mathbf{t} = \psi\mathbf{x}$. Then $S_{\mathbf{t}}^{\mathbf{x}}\mathbf{A} \in \mathcal{U}$, so by inductive hypothesis and the Substitution-Value Theorem (2301), $\mathcal{T} = \mathcal{V}^{\mathcal{M}}S_{\mathbf{t}}^{\mathbf{x}}\mathbf{A} = \mathcal{V}_{\psi}^{\mathcal{M}}\mathbf{A}$. Hence $\mathcal{M} \models \mathbf{E}$.
- (f) \mathbf{E} is $\sim \forall \mathbf{x}\mathbf{A}$. Then there is an individual constant \mathbf{c} (which is, of course, a variable-free term) such that $\sim S_{\mathbf{c}}^{\mathbf{x}}\mathbf{A} \in \mathcal{U}$, so $\mathcal{M} \models \sim S_{\mathbf{c}}^{\mathbf{x}}\mathbf{A}$ by inductive hypothesis, so $\mathcal{V}^{\mathcal{M}}S_{\mathbf{c}}^{\mathbf{x}}\mathbf{A} = \mathbf{F}$. $\models \forall \mathbf{x}\mathbf{A} \supset S_{\mathbf{c}}^{\mathbf{x}}\mathbf{A}$ by 2303, so $\mathcal{V}^{\mathcal{M}}\forall \mathbf{x}\mathbf{A} = \mathbf{F}$, so $\mathcal{M} \models \mathbf{E}$.

Since $\mathcal{S} \subseteq \mathcal{U}$, every member of \mathcal{S} is valid in \mathcal{M} , so we see that \mathcal{M} is a frugal model of \mathcal{S} . ■

REMARK. It is natural to ask whether the purity condition in Theorem 2507 is really necessary. To see that it is, consider a formulation of \mathcal{F} with k individual constants. Let p be a fixed positive integer greater than k , and let \mathcal{S} be the set of all sentences of the form $\sim \forall \mathbf{x}_1 \dots \forall \mathbf{x}_m \mathbf{B}(\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{c}_1, \dots, \mathbf{c}_n)$, where $n \geq 0$, $m + n \geq p$, $\mathbf{x}_1, \dots, \mathbf{x}_m$ are distinct variables, $\mathbf{c}_1, \dots, \mathbf{c}_n$ are distinct individual constants, and $\mathbf{B}(\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{c}_1, \dots, \mathbf{c}_n)$ is a quantifier-free wff in which $\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{c}_1, \dots, \mathbf{c}_n$ all occur. Notice that $m \geq p - k > 0$, since $m + n \geq p$ and $k \geq n$. Let Γ be the class of all subsets of \mathcal{S} .

Note that if $\sim \forall \mathbf{x}\mathbf{A} \in \mathcal{S}$, and \mathbf{c} is any individual constant which does not occur in \mathbf{A} , then $\sim S_{\mathbf{c}}^{\mathbf{x}}\mathbf{A} \in \mathcal{S}$. (It is true vacuously in the case of a sentence such as $\sim \forall \mathbf{x}_1 \mathbf{B}(\mathbf{x}_1, \mathbf{c}_1, \dots, \mathbf{c}_k)$, since there is no individual constant which does not occur in this sentence.) Therefore Γ is an abstract consistency class, and $\mathcal{S} \in \Gamma$. However, \mathcal{S} has no model, since it contains unsatisfiable sentences of the form $\sim \forall \mathbf{x}_1 \dots \forall \mathbf{x}_m [\mathbf{C} \vee \sim \mathbf{C}]$.

Note that if $p \leq k$, the counterexample fails, so when p is held fixed there is a sense in which one can make Γ fail to be an abstract consistency class by adding at least $p - k$ new individual constants to the language. However, fewer than $p - k$ will not suffice.

REMARK. For many applications of Smullyan's Unifying Principle one can simply add constants to the language without causing any essential disruption of the situation, and thus satisfy the purity condition. In such cases we shall often assume without discussion that this is done, and omit explicit mention of the purity condition.

2508 Theorem. Every consistent set of sentences has a frugal model.⁹

Proof: Let \mathcal{S} be a consistent set of sentences of \mathcal{F} . Expand \mathcal{F} by adding to it $\text{card}(\mathcal{L}(\mathcal{F}))$ individual constants; let the expanded formulation of \mathcal{F} be called \mathcal{F}^+ . \mathcal{S} is consistent in \mathcal{F}^+ by 2503. Thus \mathcal{S} has a frugal model by 2505 and 2507. ■

2509 Generalized Completeness Theorem. If \mathcal{S} is a set of sentences and \mathbf{A} is valid in all frugal models of \mathcal{S} , then $\mathcal{S} \vdash \mathbf{A}$.

Proof: Let $\overline{\mathbf{A}}$ be the universal closure of \mathbf{A} . Suppose \mathbf{A} is valid in all frugal models of \mathcal{S} . Then $\overline{\mathbf{A}}$ is too, so $\mathcal{S} \cup \{\sim \overline{\mathbf{A}}\}$ has no frugal model. Therefore $\mathcal{S} \cup \{\sim \overline{\mathbf{A}}\}$ is inconsistent by 2508, so $\mathcal{S} \vdash \sim \sim \overline{\mathbf{A}}$ by 2501, so $\mathcal{S} \vdash \mathbf{A}$ by Rule P and VI. ■

It should be noted that Theorem 2508 and the Generalized Completeness Theorem (2509) are equivalent in the sense that each can easily be derived from the other. To prove 2508 from 2509, suppose that \mathcal{S} is a set of sentences which has no frugal model. Then $[p \wedge \sim p]$ is valid in all frugal models of \mathcal{S} , so $\mathcal{S} \vdash p \wedge \sim p$ by 2509, so \mathcal{S} is inconsistent.

2510 Gödel's Completeness Theorem [Gödel, 1930]. Every valid wff of \mathcal{F} is a theorem.

Proof: Let $\mathcal{S} = \emptyset$ in 2509.

REMARK. Since we have shown that the axiom schemata and rules of inference of \mathcal{F} are independent, it is clear that each of them must be used in some way in any proof of the Completeness Theorem. The reader may wish to note where each of them has been used in the proof above.

2511 Löwenheim's Theorem [Löwenheim, 1915]. If \mathbf{A} is valid in every countable interpretation, then \mathbf{A} is valid.

Proof: We may regard \mathbf{A} as a wff of a countable formulation of \mathcal{F} . If \mathbf{A} is valid in every countable interpretation (model of \emptyset), then $\vdash \mathbf{A}$ by 2509, so $\models \mathbf{A}$ by 2303. ■

NOTE: By 2303 and 2510, the theorems of \mathcal{F} are precisely the valid wffs.

⁹This theorem was proved in [Gödel, 1930] but the proof given here (which is mostly embodied in the proof of Theorem 2507) is based on the approach introduced in [Henkin, 1949].

Thus we have two radically different characterizations of this set of wffs, one syntactic, and one semantic. The next theorem also shows the equivalence of a syntactic and a semantic notion.

2512 Theorem. A set of sentences is consistent iff it has a model.

Proof: by 2304 and 2508. ■

2513 Löwenheim-Skolem Theorem [Skolem, 1920]. Every set of sentences which has a model has a frugal model.

Proof: By 2512 and 2508. ■

The Löwenheim-Skolem theorem has a rather puzzling consequence which is known as Skolem's Paradox. Set Theory can be axiomatized within first-order logic, and this is customarily done so that one can prove the existence of infinite sets, the existence of the power set $\mathcal{P}(S)$ (collection of all subsets) of each set S , and Cantor's Theorem (which is discussed in a type-theoretic formulation in exercise X5304) that $\mathcal{P}(S)$ always has a greater cardinality than S . Of course, this means that one can prove within presumably consistent formulations of Axiomatic Set Theory the existence of uncountably infinite sets, although 2512 and 2513 imply that such a theory must have a countable model. The paradox is best understood when one has a clear notion of a non-standard model, so we postpone further discussion of it to §55 (following theorem 5503).

2514 Compactness Theorem. A set \mathcal{S} of sentences has a model iff every finite subset of \mathcal{S} has a model.

Proof: Clearly any model for \mathcal{S} is a model for each subset of \mathcal{S} . If \mathcal{S} has no model, then by 2508 \mathcal{S} is inconsistent, so if \mathbf{A} is any wff, $\mathcal{S} \vdash \mathbf{A} \wedge \sim \mathbf{A}$. Hence there is a finite subset \mathcal{H} of \mathcal{S} such that $\mathcal{H} \vdash \mathbf{A} \wedge \sim \mathbf{A}$. Thus \mathcal{H} is inconsistent, and so by 2512 has no model. ■

APPLICATION: An interesting application of the Compactness Theorem permits an extension of the celebrated Four Color Theorem [Appel and Haken, 1976] of graph theory from finite to infinite graphs.

A simple *graph* consists of a set V of *vertices* and a set of E of *edges*, which are unordered pairs of vertices. Vertices u and v are said to be *adjacent* iff $(u, v) \in E$. The graph is *finite* iff V is finite, and it is *planar* iff it can be embedded in the plane so that its edges (represented as lines) never cross.

(We refer the reader to a text on graph theory for more details.) Let C be a set of four colors, which we shall refer to as the first, second, third, and fourth color. A *4-coloring* of a graph is an assignment of colors (i.e., members of C) to vertices so that each vertex is assigned exactly one color, and adjacent vertices never receive the same color. The Four Color Theorem says that every planar graph has a 4-coloring. We shall show that once one knows that all finite planar graphs have 4-colorings, one also knows this for infinite planar graphs.

Let an infinite planar graph \mathcal{G} be given. We shall introduce a first-order theory $\mathcal{F}_\mathcal{G}$ to describe certain aspects of \mathcal{G} . For each vertex of \mathcal{G} we shall have an individual constant to denote that vertex. We shall have four monadic predicate constants C_1 , C_2 , C_3 , and C_4 , and $C_i\mathbf{a}$ will mean that vertex \mathbf{a} has color i . We shall use the notation $\bigvee_{j=1}^4 C_jx$ as an abbreviation for $[C_1x \vee C_2x \vee C_3x \vee C_4x]$, and we shall use obvious extensions of this notation.

Let \mathcal{S} be the set which contains the following sentences:

$$\begin{aligned} & \forall x \bigvee_{j=1}^4 C_jx \\ & \forall x \bigwedge_{1 \leq j < k \leq 4} \sim [C_jx \wedge C_kx] \\ & \bigwedge_{j=1}^4 \sim [C_j\mathbf{a} \wedge C_j\mathbf{b}], \end{aligned}$$

where \mathbf{a} and \mathbf{b} are any constants denoting adjacent vertices.

In the intended interpretation, the vertices of \mathcal{G} are the individuals and the sentences of \mathcal{S} express the statements that every vertex has a color, that every vertex has only one color, and that adjacent vertices do not have the same color. Of course, \mathcal{S} will be infinite if \mathcal{G} has infinitely many edges.

Let \mathcal{S}' be any finite subset of \mathcal{S} , and let \mathcal{G}' be the subgraph of \mathcal{G} which contains those vertices whose names occur in sentences of \mathcal{S}' . Since \mathcal{G}' is finite and planar, it has a 4-coloring, and if one interprets the constants of $\mathcal{F}_\mathcal{G}$ as indicated above with respect to this coloring, the sentences of \mathcal{S}' will be true in the interpretation where the domain of individuals is the set of vertices of \mathcal{G}' . Thus every finite subset of \mathcal{S} has a model, so by the Compactness Theorem, \mathcal{S} has a model. If one assigns colors to vertices so that a vertex with name \mathbf{a} is given the unique color j such that $C_j\mathbf{a}$ is true in this model, one obtains a 4-coloring of the entire graph \mathcal{G} .

APPLICATION: Another important application of the Compactness Theorem leads to nonstandard number theory and nonstandard analysis. Let \mathcal{T} be a first-order theory which is a formalization of the theory of natural numbers. Thus, \mathcal{T} may have constants denoting certain natural numbers, certain functions of natural numbers (such as successor, addition, and multiplication), and certain properties of natural numbers and relations between natural numbers. Also, \mathcal{T} will have postulates (perhaps Peano's Postulates¹⁰ and definitions of the functions and relations denoted by constants of \mathcal{T}) which are true in a model \mathcal{N} which has the set of natural numbers as its domain of individuals. Let us assume that each natural number n is denoted by some variable-free term \bar{n} , and that the equality relation is definable by some wff of \mathcal{T} . Let \mathcal{T}^* be the expansion and extension of \mathcal{T} obtained by adding a new individual constant c and the set of postulates $\{\sim c = \bar{n} \mid n \text{ is a natural number}\}$. It is easy to see that every finite subset of the postulates of \mathcal{T}^* has a model (simply use \mathcal{N} with an appropriate interpretation for c), so by the Compactness Theorem \mathcal{T}^* has a model \mathcal{N}^* , which is called a *nonstandard model for number theory*. Since the denotation of c in \mathcal{N}^* is not any finite number, \mathcal{N}^* may be said to contain "infinite" numbers. Of course, \mathcal{N}^* is a model of \mathcal{T} , so we see that all formalizations of number theory have nonstandard models.

We may even abstractly consider a theory \mathcal{T} which has constants to denote all functions and relations on the set of natural numbers, whose postulates are all the sentences of \mathcal{T} which are true in \mathcal{N} . (It will be seen that such a set of postulates cannot be effectively specified, but that is irrelevant for our present purposes.) Then every true statement about natural numbers expressible in \mathcal{T} will also be true in the nonstandard model \mathcal{N}^* . This rather paradoxical situation may be regarded as revealing an inadequacy in the expressive power of first order logic.

Similar considerations apply to formalizations of other infinite mathematical systems. For another example, let \mathcal{T} be a first-order theory for the real numbers, and let \mathcal{T}^* be obtained by adding to \mathcal{T} a new constant c and the set of postulates $\{0 < c \wedge c < \bar{r} \mid r \text{ is a positive real number}\}$. \mathcal{T}^* will have a model in which there are numbers, called *infinitesimals*, which are smaller than all standard real numbers, but greater than zero.

¹⁰Properly speaking, the Induction Axiom cannot be expressed in first-order logic (unless one uses some device such as embedding two-sorted logic or axiomatic set theory into first-order logic). Nevertheless, its first-order consequences can be expressed by the infinite set of wffs of the form $[A(\bar{0}) \wedge \forall x [A(x) \supset A(\bar{s}x)]] \supset \forall x A(x)$, where $\bar{0}$ is a constant denoting the number zero, \bar{s} is a function constant denoting the successor function, and $A(x)$ is a wff in which x is free.

We can now easily extend 2503 by using a semantic, rather than syntactic, proof.

2515 Proposition. Let \mathcal{F}^1 and \mathcal{F}^2 be formulations of \mathcal{F} , so that \mathcal{F}^2 is an expansion of \mathcal{F}^1 obtained by adding additional constants (of any type) to the set of constants of \mathcal{F}^1 . Let \mathcal{G} be any set of sentences of \mathcal{F}^1 . Then

- (1) $\mathcal{F}^2 \cup \mathcal{G}$ is a conservative extension of $\mathcal{F}^1 \cup \mathcal{G}$.
- (2) \mathcal{G} is consistent in \mathcal{F}^1 iff it is consistent in \mathcal{F}^2 .

Proof: Let \mathbf{C} be any wff of \mathcal{F}^1 . Clearly if $\vdash_{\mathcal{F}^1 \cup \mathcal{G}} \mathbf{C}$, then $\vdash_{\mathcal{F}^2 \cup \mathcal{G}} \mathbf{C}$. Suppose $\vdash_{\mathcal{F}^2 \cup \mathcal{G}} \mathbf{C}$, and let \mathcal{M} be any model of $\mathcal{F}^1 \cup \mathcal{G}$. \mathcal{M} can be extended to a model \mathcal{M}' of $\mathcal{F}^2 \cup \mathcal{G}$ simply by providing arbitrary interpretations for the new constants of \mathcal{F}^2 . $\mathcal{M}' \models \mathbf{C}$ by 2303, so $\mathcal{M} \models \mathbf{C}$, since the truth value of a wff is not affected by constants which do not occur in it. Thus \mathbf{C} is valid in all models of $\mathcal{F}^1 \cup \mathcal{G}$, so $\vdash_{\mathcal{F}^1 \cup \mathcal{G}} \mathbf{C}$ by 2509. This proves (1).

(2) follows from (1) and 2502. ■

§25A. Supplement: Simplified Completeness Proof

In this section we provide a direct proof of Theorem 2508 for countable formulations of \mathcal{F} , bypassing Theorems 2504-2507 and the associated definitions.

2508A Theorem. Every consistent set of sentences of a countable formulation of \mathcal{F} has a countable model.

Proof: Let \mathcal{S} be a consistent set of sentences of \mathcal{F} . The theorem is trivial if \mathcal{S} is empty, so we may assume \mathcal{S} is non-empty. Expand \mathcal{F} by adding to it countably many individual constants; let the expanded formulation of \mathcal{F} be called \mathcal{F}^+ . Enumerate the sentences of \mathcal{F}^+ , and for each index τ , let \mathbf{S}_τ be the τ -th sentence in this enumeration.

Let \mathcal{C} be a countable set of individual constants which do not occur in sentences of \mathcal{S} . By induction, for each index τ we define a set \mathcal{S}_τ of sentences so that $\mathcal{S}_\sigma \subseteq \mathcal{S}_\tau$ whenever $\sigma \leq \tau$, and for each τ only finitely many constants from \mathcal{C} occur in the sentences of \mathcal{S}_τ . $\mathcal{S}_0 = \mathcal{S}$. The definition for successor indices $\tau + 1$ is by cases:

- (1) If $\mathcal{S}_\tau \cup \{\mathbf{S}_{\tau+1}\}$ is not consistent, then $\mathcal{S}_{\tau+1} = \mathcal{S}_\tau$.

- (2) If $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\}$ is consistent, and \mathbf{S}_τ is not of the form $\sim \forall x A$, then $\mathcal{S}_{\tau+1} = \mathcal{S}_\tau \cup \{\mathbf{S}_\tau\}$.
- (3) If $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\}$ is consistent and \mathbf{S}_τ is $\sim \forall x A$, then $\mathcal{S}_{\tau+1} = \mathcal{S}_\tau \cup \{\sim \forall x A, \sim S_c^x A\}$, where c is the first individual constant from C which does not occur in any sentence of $\mathcal{S}_\tau \cup \{\mathbf{S}_\tau\}$. (Note that there must be such a constant.)

It is readily seen by induction that only finitely many constants from C occur in the sentences of each \mathcal{S}_τ .

We next show by induction that \mathcal{S}_τ is consistent for each τ . $\mathcal{S}_0 = \mathcal{S}$, which is consistent in \mathcal{F}^+ by Theorem 2503. It is clear that if \mathcal{S}_τ is consistent, then $\mathcal{S}_{\tau+1}$ is consistent in cases (1) and (2); we must show that this is also true in case (3). Suppose $\mathcal{S}_\tau \cup \{\sim \forall x A\}$ is consistent but $\mathcal{S}_\tau \cup \{\sim \forall x A, \sim S_c^x A\}$ is inconsistent, where c is an individual constant which does not occur in \mathcal{S}_τ or A . Then $\mathcal{S}_\tau \cup \{\sim \forall x A\} \vdash S_c^x A$ by Theorem 2501 and Rule P, so there is a finite subset \mathcal{H} of $\mathcal{S}_\tau \cup \{\sim \forall x A\}$ such that $\mathcal{H} \vdash S_c^x A$. Given a proof \mathcal{P} of $S_c^x A$ from \mathcal{H} , let y be an individual variable which does not occur in \mathcal{P} or in \mathcal{H} . Since c does not occur in \mathcal{H} or A , it can be seen that $S_y^c \mathcal{P}$ is a proof of $S_y^x A$ from \mathcal{H} , so $\mathcal{H} \vdash S_y^x A$. Hence, $\mathcal{H} \vdash \forall y S_y^x A$ by Generalization, and $\mathcal{H} \vdash \forall x A$ by the α Rule, so $\mathcal{S}_\tau \cup \{\sim \forall x A\} \vdash \forall x A$, so $\mathcal{S}_\tau \cup \{\sim \forall x A\}$ is inconsistent. This contradiction proves our assertion.

Let $\mathcal{U} = \bigcup_{\tau < \omega} \mathcal{S}_\tau$. Clearly \mathcal{U} is a set of sentences and $\mathcal{S} \subseteq \mathcal{U}$. We show that \mathcal{U} is consistent. Let \mathcal{H} be any finite subset of \mathcal{U} . Then there is a τ such that $\mathcal{H} \subseteq \mathcal{S}_\tau$, and \mathcal{S}_τ is consistent, so \mathcal{H} is also consistent. Thus every finite subset of \mathcal{U} is consistent, so \mathcal{U} is consistent.

Note that (*) if D is a sentence such that $\mathcal{U} \cup \{D\}$ is consistent, then $D \in \mathcal{U}$. For there is an index τ such that $\mathbf{S}_\tau = D$, and $\mathcal{S}_\tau \cup \{D\} \subseteq \mathcal{U} \cup \{D\}$, which is consistent; thus $\mathcal{S}_\tau \cup \{D\}$ is consistent. Hence $D \in \mathcal{S}_{\tau+1} \subseteq \mathcal{U}$, so $D \in \mathcal{U}$. From (*) it follows that (**) if $\mathcal{U} \vdash D$, then $D \in \mathcal{U}$.

We assert that \mathcal{U} has the following properties for all wffs A and B :

- (a) There is no atomic sentence A such that $A \in \mathcal{U}$ and $\sim A \in \mathcal{U}$.
- (b) If $\sim \sim A \in \mathcal{U}$, then $A \in \mathcal{U}$.
- (c) If $[A \vee B] \in \mathcal{U}$, then $A \in \mathcal{U}$ or $B \in \mathcal{U}$.
- (d) If $\sim [A \vee B] \in \mathcal{U}$, then $\sim A \in \mathcal{U}$ and $\sim B \in \mathcal{U}$.
- (e) If $\forall x A \in \mathcal{U}$, then $S_t^x A \in \mathcal{U}$ for each variable-free term t .
- (f) If $\sim \forall x A \in \mathcal{U}$, then there is an individual constant c such that $\sim S_c^x A \in \mathcal{U}$.

(a) is true since \mathcal{U} is consistent, (b), (d) and (e) follow from (**). To prove (c), suppose $[A \vee B] \in \mathcal{U}$. Then $\mathcal{U} \cup \{A\}$ is consistent or $\mathcal{U} \cup \{B\}$ is consistent, so $A \in \mathcal{U}$ or $B \in \mathcal{U}$ by (*). To prove (f), suppose $\sim \forall x A \in \mathcal{U}$. Choose τ such that $S_\tau = \sim \forall x A$. Then $S_\tau \cup \{\sim \forall x A\} \subseteq \mathcal{U}$, which is consistent, so $S_\tau \cup \{\sim \forall x A\}$ is consistent. Thus $S_{\tau+1} = S_\tau \cup \{\sim \forall x A, \sim S_c^x A\}$ for some individual constant c , so $\sim S_c^x A \in S_{\tau+1} \subseteq \mathcal{U}$.

We next define an interpretation $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$. Let \mathcal{D} be the set of variable-free terms of \mathcal{F}^+ . Clearly \mathcal{D} is countable. For each function constant f^n , if t_1, \dots, t_n are in \mathcal{D} , let $(\mathcal{J}f^n)t_1, \dots, t_n = f^n t_1 \dots t_n$, and if a is an individual constant, let $\mathcal{J}a = a$. (Thus variable-free terms denote themselves.) For each predicate constant P^n of \mathcal{F}^+ , if t_1, \dots, t_n are in \mathcal{D} , let $(\mathcal{J}P^n)t_1 \dots t_n = T$ iff $P^n t_1 \dots t_n \in \mathcal{U}$. Similarly, if p is a propositional constant, $\mathcal{J}p = T$ iff $p \in \mathcal{U}$.

Clearly $\mathcal{V}^\mathcal{M}t = t$ for each variable-free term t , and for each atomic sentence E , $\mathcal{M} \models E$ iff $E \in \mathcal{U}$.

We show by induction on the number of occurrences of \sim , \vee , and \forall in E that if $E \in \mathcal{U}$, then $\mathcal{M} \models E$. We suppose that $E \in \mathcal{U}$, and consider the following cases:

- (a) E is $\sim A$ for some atomic sentence A . Since $E \in \mathcal{U}$, A cannot be in \mathcal{U} , so $\mathcal{V}^\mathcal{M}A = F$. Hence $\mathcal{M} \models E$.
- (b) E is $\sim\sim A$. Then $A \in \mathcal{U}$, so $\mathcal{M} \models A$ by inductive hypothesis, so $\mathcal{M} \models E$.
- (c) E is $[A \vee B]$. Then $A \in \mathcal{U}$ or $B \in \mathcal{U}$, so by inductive hypothesis $\mathcal{M} \models A$ or $\mathcal{M} \models B$. Hence $\mathcal{M} \models E$ in either case.
- (d) E is $\sim[A \vee B]$. Then $\sim A \in \mathcal{U}$ and $\sim B \in \mathcal{U}$ so $\mathcal{M} \models \sim A$ and $\mathcal{M} \models \sim B$ by inductive hypothesis, so $\mathcal{V}^\mathcal{M}A = F = \mathcal{V}^\mathcal{M}B$, so $\mathcal{V}^\mathcal{M}[A \vee B] = F$ and $\mathcal{M} \models E$.
- (e) E is $\forall x A$. Let ψ be any assignment, and let $t = \psi x$. Then $S_t^x A \in \mathcal{U}$, so by inductive hypothesis and the Substitution-Value Theorem (2302), $T = \mathcal{V}^\mathcal{M}S_t^x A = \mathcal{V}_\psi^\mathcal{M}A$. Hence $\mathcal{M} \models E$.
- (f) E is $\sim \forall x A$. Then there is an individual constant c (which is, of course, a variable-free term) such that $\sim S_c^x A \in \mathcal{U}$, so $\mathcal{M} \models \sim S_c^x A$ by inductive hypothesis, so $\mathcal{V}^\mathcal{M}S_c^x A = F$. $\models \forall x A \supset S_c^x A$ by 2303, so $\mathcal{V}^\mathcal{M}\forall x A = F$, so $\mathcal{M} \models E$.

Since $\mathcal{S} \subseteq \mathcal{U}$, every member of \mathcal{S} is valid in \mathcal{M} , so we see that \mathcal{M} is a countable model of \mathcal{S} . ■

EXERCISES

X2500. In the proof of Theorem 2507, suppose that Γ (and therefore Δ) is the class of consistent sets of sentences. Show that in this case for each sentence \mathbf{E} :

- (1) $\mathbf{E} \in \mathcal{U}$ or $\sim \mathbf{E} \in \mathcal{U}$.
- (2) $\mathbf{E} \in \mathcal{U}$ if and only if $\mathcal{M} \models \mathbf{E}$.

X2501. Let Γ be the class of all finite sets of quantifier-free sentences whose conjunctions are not contradictory (i.e., not substitution instances of contradictions of propositional calculus). Show that Γ is an abstract consistency class.

X2502. Show how to derive Löwenheim's Theorem (2511) from the Löwenheim-Skolem Theorem (2513) without using any other theorems from §25.

X2503. A set \mathcal{H} of sentences of \mathcal{F} is said to be *complete* if for every sentence \mathbf{A} of \mathcal{F} , either $\mathcal{H} \vdash \mathbf{A}$ or $\mathcal{H} \vdash \sim \mathbf{A}$. (Do not confuse this use of the word with other uses.) Show that if \mathcal{G} is any consistent set of sentences, then there is a complete, consistent set \mathcal{H} of sentences such that $\mathcal{G} \subseteq \mathcal{H}$.

X2504. We say that \mathbf{B} is a (*semantic*) *consequence* of \mathbf{A} if for every interpretation \mathcal{M} and every assignment ϕ , if $\mathcal{M} \models_{\phi} \mathbf{A}$ then $\mathcal{M} \models_{\phi} \mathbf{B}$. Consider the following assertions:

- (1) If \mathbf{B} is a consequence of \mathbf{A} , then \mathbf{B} is valid in all models of \mathbf{A} .
- (2) If \mathbf{B} is valid in all models of \mathbf{A} , then \mathbf{B} is a consequence of \mathbf{A} .

Are these assertions true for all wffs \mathbf{A} and \mathbf{B} ? If so, prove them; if not, provide counterexamples.

X2505. Show how Theorem 2514 can be used to prove that every partial ordering of a set can be extended to a total ordering of the set, once one knows that this is true for finite sets.

§26. Equality

The system $\mathcal{F}^=$ of *first-order logic with equality* is obtained by adding to the primitive symbols of \mathcal{F} a binary predicate constant $=$ (unless this constant

is already present)¹¹ and adding to the axioms of \mathcal{F} the following axioms (where $s = t$ is an abbreviation for $= st$):

Axiom 6. $x = x$ (Reflexivity)

Axiom Schema 7. $x = y \supset \neg S_x^z \supset S_y^z A$ (Substitutivity)
where A is an atomic wff.

A first order theory is a *first-order theory with equality* if it has a binary predicate $=$ such that the wffs above are theorems of the theory.

It can be seen (Exercise X2601) that the elementary theory \mathcal{G} of additive abelian groups which was presented in §21 is an example of such a theory. It can also be seen (Exercise X2602) that the notion of equality is unambiguous in such a theory.

We shall often write $s \neq t$ as an abbreviation for $\sim s = t$.

In any first order theory with equality, the following are theorems:

2600. $\vdash x = y \supset y = x$

Proof:

- (.1) $\vdash x = y \supset \neg x = x \supset y = x$ Axiom 7
(.2) $\vdash x = y \supset y = x$ Rule P: .1, Axiom 6

■

2601. $\vdash x = y \supset \neg y = z \supset x = z$

Proof:

- (.1) $\vdash y = x \supset \neg y = z \supset x = z$ Axiom 7
(.2) $\vdash x = y \supset \neg y = z \supset x = z$ Rule P: .1, 2600

■

2602. $\vdash x = y \supset \neg (S_x^z A) \equiv S_y^z A$, where x and y are free for z in A .

Proof: by induction on the construction of A

Case 1. A is atomic.

- (.1) $\vdash x = y \supset \neg S_x^z A \supset S_y^z A$ Axiom 7

¹¹Actually, it is not essential that $=$ be a primitive symbol, since in some first-order theories $=$ can be defined. For example, in Axiomatic Set Theory $x = y$ is sometimes regarded as an abbreviation for $\forall z [z \in x \equiv z \in y]$ or for $\forall w [x \in w \equiv y \in w]$.

- (.2) $\vdash y = x \supset \neg S_y^z A \supset S_x^z A$ Axiom 7
(.3) $\vdash x = y \supset y = x$ Sub: 2600
(.4) $\vdash x = y \supset \neg (S_x^z A) \equiv S_y^z A$ Rule P: .1, .2, .3

The cases where A has the form $\sim B$ or $[B \vee C]$ follow directly by Rule P from the inductive hypotheses.

Case 2. A has the form $\forall w B$. If z is not free in A , this instance of 2602 follows by Rule P, so we may assume z is free in A . Thus, since x and y are free for z in A , they are distinct from w and free for z in B .

- $$\begin{aligned} \vdash x = y \supset \neg (S_x^z B) \equiv S_y^z B & \quad \text{by inductive hypothesis} \\ \vdash x = y \supset \forall w \neg (S_x^z B) \equiv S_y^z B & \quad \supset \forall \text{ Rule (2103)} \\ \vdash x = y \supset \neg \forall w (S_x^z B) \equiv \forall w S_y^z B & \quad \text{Rule P: 2107} \end{aligned}$$
-

2603. $\vdash x = y \supset (S_x^z t) = (S_y^z t)$ for any term t .

Proof: Let w be an individual variable distinct from x, y, z and the variables in t .

- $$\begin{aligned} (.1) \vdash x = y \supset \neg w = (S_x^z t) \supset w = S_y^z t & \quad \text{Axiom 7 } (A \text{ is } w = t) \\ (.2) \vdash x = y \supset \neg (S_x^z t) = (S_x^z t) \supset (S_x^z t) = S_y^z t & \quad \text{Sub: .1} \\ (.3) \vdash x = y \supset (S_x^z t) = S_y^z t & \quad \text{Sub: Axiom 6; Rule P: .2} \end{aligned}$$
-

2604. $\vdash x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset f x_1 \dots x_n = f y_1 \dots y_n$ for any n -ary function symbol f .

Proof:

$$(a) \vdash x_1 = y_1 \supset f x_1 x_2 \dots x_n = f y_1 x_2 \dots x_n \quad 2603$$

We show by induction on i that for $1 \leq i \leq n$,

$$(b_i) \vdash x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset f x_1 \dots x_n = f y_1 \dots y_i x_{i+1} \dots x_n$$

The basis (b_1) of the induction follows from (a) by Rule P. For the induction step we assume (b_i) as inductive hypothesis.

$$(c) \vdash x_{i+1} = y_{i+1} \supset f y_1 \dots y_i x_{i+1} x_{i+2} \dots x_n = f y_1 \dots y_i y_{i+1} x_{i+2} \dots x_n \quad 2603$$

$$(b_{i+1}) \vdash x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \mathbf{f}x_1 \dots x_n = \mathbf{f}y_1 \dots y_{i+1}x_{i+2} \dots x_n$$

Sub: 2601; Rule P: b_i, c

This completes the inductive proof. (b_n) is 2604. ■

2605. $\vdash x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \blacksquare \mathbf{P}x_1 \dots x_n \equiv \mathbf{P}y_1 \dots y_n$ for any n -ary predicate symbol \mathbf{P} .

The proof is similar to that of 2604.

2606 Proposition. Let \mathcal{T} be a first-order theory with a binary predicate constant $=$ such that

$$(a) \vdash_{\mathcal{T}} x = x$$

and for each n -ary function symbol \mathbf{f}^n

$$(b) \vdash_{\mathcal{T}} x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \mathbf{f}^n x_1 \dots x_n = \mathbf{f}^n y_1 \dots y_n$$

and for each n -ary predicate symbol \mathbf{P}^n

$$(c) \vdash_{\mathcal{T}} x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \blacksquare \mathbf{P}_1^n \dots x_n \supset \mathbf{P}^n y_1 \dots y_n.$$

Then \mathcal{T} is a first-order theory with equality.

Proof: In this proof we shall let $\vdash \mathbf{A}$ mean $\vdash_{\mathcal{T}} \mathbf{A}$.

We first show by induction on the construction of t that if t is any term, then

(d) $\vdash x = y \supset (S_x^z t) = S_y^z t$. If t is an individual constant or variable, then (d) is easily proved using (a), Sub, and Rule P. Suppose t is $f^n t_1 \dots t_n$. For $1 \leq i \leq n$, let t'_i be $S_x^z t_i$ and t''_i be $S_y^z t_i$. Then

$$(e_i) \vdash x = y \supset t'_i = t''_i \quad \text{by inductive hypothesis}$$

$$(f) \vdash t'_1 = t''_1 \wedge \dots \wedge t'_n = t''_n \supset f^n t'_1 \dots t'_n = f^n t''_1 \dots t''_n \quad \text{Sub: (b)}$$

$$(g) \vdash x = y \supset f^n t'_1 \dots t'_n = f^n t''_1 \dots t''_n \quad \text{Rule P: (e₁)-(e_n), (f)}$$

This completes the inductive proof of (d).

Now let \mathbf{A} be an atomic wff $P^n t_1 \dots t_n$.

$$(h_i) \vdash x = y \supset (S_x^z t_i) = S_y^z t_i \quad (d)$$

$$(j) \vdash x = y \supset \blacksquare S_x^z A \supset S_y^z A \quad \text{Sub: (c); Rule P: (h₁)-(h_n)}$$

Since (j) is Axiom Schema 7, this completes the proof. ■

DEFINITION. In any first order theory with equality,

$$\exists_1 \mathbf{x} \mathbf{A} \text{ stands for } \exists \mathbf{x} \mathbf{A} \wedge \forall \mathbf{x} \forall \mathbf{y} [\mathbf{A} \wedge (\mathbf{S}_{\mathbf{y}}^{\mathbf{x}} \mathbf{A}) \supset \mathbf{x} = \mathbf{y}],$$

where \mathbf{y} is the first individual variable distinct from \mathbf{x} and all variables which occur in \mathbf{A} . Clearly, $\exists_1 \mathbf{x} \mathbf{A}$ means “there exists a unique \mathbf{x} such that \mathbf{A} ”. Alternative possible definitions are suggested by the following theorems of any first-order theory with equality:

$$2607. \vdash \exists_1 x Px \equiv \exists x \bullet Px \wedge \forall y \bullet Py \supset x = y$$

$$2608. \vdash \exists_1 x Px \equiv \exists y \forall x \bullet Px \equiv x = y$$

The proofs of 2607 and 2608 are left to the reader.

It follows from Axiom 6, 2600, and 2601 that in any model for a first-order theory with equality, the interpretation of the equality predicate must be an equivalence relation on the domain of individuals. We next see that one can actually require this relation to the identity relation.

DEFINITION. An interpretation [model] $\langle \mathcal{D}, \mathcal{J} \rangle$ for a first-order language with an equality predicate $=$ is an *equality-interpretation* [model] iff $\mathcal{J}=$ is the identity relation on \mathcal{D} , i.e., for all elements d and e of \mathcal{D} , $(\mathcal{J}=)de$ is T iff d is e .

2609 Soundness Theorem for Logic with Equality. If \mathcal{M} is an equality model for a set \mathcal{S} of sentences and $\mathcal{S} \vdash_{\mathcal{F}=} \mathbf{A}$, then $\mathcal{M} \models \mathbf{A}$.

Proof: Let \mathcal{F} be the formulation of first-order logic with $\mathcal{L}(\mathcal{F}) = \mathcal{L}(\mathcal{F}^=)$, and let \mathcal{E} be the set of universal closures of Axioms 6 and 7. We first show that any equality-model \mathcal{M} for \mathcal{S} is a model for \mathcal{E} .

Clearly $\mathcal{M} \models$ Axiom 6. To treat Axiom 7, let ϕ be an assignment into \mathcal{M} . If $\varphi \mathbf{x} \neq \varphi \mathbf{y}$, then $\mathcal{V}_\varphi[\mathbf{x} = \mathbf{y}] = F$, so $\mathcal{M} \models_\varphi$ Axiom 7. If $\varphi \mathbf{x} = \varphi \mathbf{y}$, let ψ be that assignment which agrees with φ off \mathbf{z} , while $\psi \mathbf{z} = \varphi \mathbf{x} = \varphi \mathbf{y}$. Then by the Substitution-Value Theorem (2302), $\mathcal{V}_\varphi \mathbf{S}_{\mathbf{x}}^{\mathbf{z}} \mathbf{A} = \mathcal{V}_\psi \mathbf{A} = \mathcal{V}_\psi \mathbf{S}_{\mathbf{y}}^{\mathbf{z}} \mathbf{A}$, so $\mathcal{M} \models_\varphi$ Axiom 7.

Thus \mathcal{M} is a model for $\mathcal{S} \cup \mathcal{E}$. Since $\mathcal{S} \vdash_{\mathcal{F}=} \mathbf{A}$, we have $\mathcal{S} \cup \mathcal{E} \vdash_{\mathcal{F}} \mathbf{A}$, so $\mathcal{M} \models \mathbf{A}$ by the Soundness Theorem for first-order logic (2303). ■

DEFINITION. Interpretations \mathcal{M} and \mathcal{N} of \mathcal{F} are *elementarily equivalent* iff for every sentence \mathbf{A} of \mathcal{F} , $\mathcal{M} \models \mathbf{A}$ iff $\mathcal{N} \models \mathbf{A}$.

2610 Proposition. Let \mathcal{M} be any model for Axioms 6 and 7. Then there is an equality-model \mathcal{N} for Axioms 6 and 7 such that the cardinality of \mathcal{N} is \leq the cardinality of \mathcal{M} , and \mathcal{N} is elementarily equivalent to \mathcal{M} .

Proof: Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$. Let \mathcal{E} be the set of equivalence classes of \mathcal{D} under the equivalence relation $(\mathcal{I} =)$. Define an interpretation function \mathcal{J} into \mathcal{E} as follows. If a is an individual constant, $\mathcal{J}a$ is the equivalence class $[\mathcal{I}a]$ containing $\mathcal{I}a$. If f^n is any n -ary function constant, P^n is any n -ary predicate constant, and d_1, \dots, d_n are elements of \mathcal{D} , then $(\mathcal{J}f^n)[d_1] \dots [d_n]$ is $[(\mathcal{I}f^n)d_1 \dots d_n]$ and $(\mathcal{J}P^n)[d_1] \dots [d_n]$ is $(\mathcal{I}P^n)d_1 \dots d_n$. It is easily seen that this definition is independent of the particular representative d_i chosen from the equivalence class $[d_i]$, since $\mathcal{M} \models 2604$ and $\mathcal{M} \models 2605$. Hence if $[d_i] = [e_i]$ for $1 \leq i \leq n$, then $[(\mathcal{I}f^n)d_1 \dots d_n]$ is $[(\mathcal{I}f^n)e_1 \dots e_n]$ and $(\mathcal{I}P^n)d_1 \dots d_n$ is $(\mathcal{I}P^n)e_1 \dots e_n$, so the definition of \mathcal{J} is unambiguous.

Let $\mathcal{N} = \langle \mathcal{E}, \mathcal{J} \rangle$. Then \mathcal{N} is an equality-model, since $(\mathcal{J} =)[d][e]$ is T iff $(\mathcal{I} =)de$ is T iff $[d]$ is $[e]$. Clearly the cardinality of \mathcal{E} is \leq that of \mathcal{D} .

For each assignment φ into \mathcal{M} , let $*\varphi$ be that assignment into \mathcal{N} such that $(* \varphi)x = [\varphi x]$ for each individual variable x . It is easily seen by induction on the construction of t that for any term t and assignment φ into \mathcal{M} , $\mathcal{V}_{*\varphi}^{\mathcal{N}}t = [\mathcal{V}_{\varphi}^{\mathcal{M}}t]$. We show by induction on the construction of A that $\mathcal{V}_{*\varphi}^{\mathcal{N}}A = \mathcal{V}_{\varphi}^{\mathcal{M}}A$ for any assignment φ into \mathcal{M} and any wff A . The cases where A is atomic or of the form $\sim B$ or $[B \vee C]$ are routine, and details are left to the reader.

Case. A is $\forall x B$.

Suppose $\mathcal{V}_{\varphi}^{\mathcal{M}}A = F$. Hence there is an assignment ψ into \mathcal{M} which agrees with φ off x such that $\mathcal{V}_{\psi}^{\mathcal{M}}B = F$. For any variable y distinct from x , $(* \psi)y = [\psi y] = [\varphi y] = (* \varphi)y$, so $*\psi$ agrees with $*\varphi$ off x . By inductive hypothesis $\mathcal{V}_{*\psi}^{\mathcal{N}}B = \mathcal{V}_{\psi}^{\mathcal{M}}B = F$, so $\mathcal{V}_{*\varphi}^{\mathcal{N}}A = F$.

Suppose $\mathcal{V}_{*\varphi}^{\mathcal{N}}A = F$. Then there is an assignment τ into \mathcal{N} which agrees with $*\varphi$ off x such that $\mathcal{V}_{\tau}^{\mathcal{N}}B = F$. Let ψ be an assignment into \mathcal{M} which agrees with φ off x such that ψx is a member of the equivalence class τx . Then $(* \psi)x = [\psi x] = \tau x$, and for any variable y distinct from x , $(* \psi)y = [\psi y] = [\varphi y] = (* \varphi)y = \tau y$, so $*\psi = \tau$. By inductive hypothesis $\mathcal{V}_{\psi}^{\mathcal{M}}B = \mathcal{V}_{*\psi}^{\mathcal{N}}B = \mathcal{V}_{\tau}^{\mathcal{N}}B = F$, so $\mathcal{V}_{\varphi}^{\mathcal{M}}A = F$.

This completes the inductive proof. It immediately follows that $\mathcal{N} \models A$ iff $\mathcal{M} \models A$ for any sentence A . ■

2611 Theorem. Let \mathcal{G} be a set of sentences consistent in $\mathcal{F}^=$. Then \mathcal{G} has a frugal equality-model.

Proof: Let \mathcal{E} be the set of universal closures of Axioms 6 and 7 of $\mathcal{F}^=$. Clearly $\mathcal{G} \cup \mathcal{E}$ is consistent in first-order logic, so $\mathcal{G} \cup \mathcal{E}$ has a frugal model by 2508. This is a model for Axioms 6 and 7, so by 2610 one obtains a frugal equality-model for \mathcal{G} . ■

2612 Extended Completeness and Soundness Theorem. Let \mathcal{S} be a set of sentences and \mathbf{A} be a wff of $\mathcal{F}^=$. Then $\mathcal{S} \vdash_{\mathcal{F}^=} \mathbf{A}$ iff \mathbf{A} is valid in all frugal equality-models of \mathcal{S} .

Proof: Let $\overline{\mathbf{A}}$ be the universal closure of \mathbf{A} . If not $\mathcal{S} \vdash_{\mathcal{F}^=} \mathbf{A}$, then $\mathcal{S} \cup \{\sim \overline{\mathbf{A}}\}$ is consistent in $\mathcal{F}^=$, and has a frugal equality-model by 2611, so \mathbf{A} is not valid in all frugal equality-models of \mathcal{S} .

Theorem 2609 completes the argument. ■

2613 Löwenheim-Skolem-Tarski Theorem. If \mathcal{S} is a set of sentences of $\mathcal{F}^=$ which has an infinite equality-model, and K is any cardinal number such that $K \geq \text{card}(\mathcal{L}(\mathcal{F}^=))$, then \mathcal{S} has an equality-model of cardinality K .

Proof: Let \mathcal{M} be an infinite equality-model for \mathcal{S} . Let \mathcal{C} be a set of new individual constants of cardinality K , and let $\mathcal{F}^{=+}$ be the expansion of $\mathcal{F}^=$ obtained by adding these constants to the primitive symbols of $\mathcal{F}^=$. Let \mathcal{H} be the set of all sentences of the form $\mathbf{c} \neq \mathbf{d}$, where \mathbf{c} and \mathbf{d} are distinct constants in \mathcal{C} . Every finite subset of $\mathcal{S} \cup \mathcal{H}$ has an equality-model, since \mathcal{M} is infinite and the finitely many constants of \mathcal{C} which actually occur in the sentences of this finite subset can be interpreted as distinct individuals of \mathcal{M} . Hence every finite subset of $\mathcal{S} \cup \mathcal{H}$ is consistent (in $\mathcal{F}^{=+}$), so $\mathcal{S} \cup \mathcal{H}$ is consistent. Thus by 2611, $\mathcal{S} \cup \mathcal{H}$ has an equality-model \mathcal{N} of cardinality $\leq K$. Since every sentence of \mathcal{H} is true in \mathcal{N} , the cardinality of \mathcal{N} cannot be less than K , so it is precisely K . ■

We next show that the problem of provability in $\mathcal{F}^=$ can be reduced to the problem of provability in \mathcal{F} .

2614 Proposition. Suppose \mathcal{F} is a formulation of first-order logic such that $\mathcal{L}(\mathcal{F}) = \mathcal{L}(\mathcal{F}^=)$. For any wff \mathbf{D} of \mathcal{F} , let \mathbf{B} be the conjunction of all wffs of the form

$$\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n [x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \mathbf{f}x_1 \dots x_n = \mathbf{f}y_1 \dots y_n],$$

where \mathbf{f} is a function symbol which occurs in \mathbf{D} , and let \mathbf{C} be the conjunction of all wffs of the form

$$\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n [x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \\ \bullet \mathbf{P} x_1 \dots x_n \supset \mathbf{P} y_1 \dots y_n],$$

where \mathbf{P} is $=$ or a predicate symbol which occurs in \mathbf{D} . Then $\vdash_{\mathcal{F}^=} \mathbf{D}$ iff $\vdash_{\mathcal{F}} \forall x [x = x] \wedge \mathbf{B} \wedge \mathbf{C} \supset \mathbf{D}$.

Proof: Suppose $\vdash_{\mathcal{F}} \forall x [x = x] \wedge \mathbf{B} \wedge \mathbf{C} \supset \mathbf{D}$. Then this wff is also a theorem of $\mathcal{F}^=$, and $\vdash_{\mathcal{F}^=} \forall x [x = x] \wedge \mathbf{B} \wedge \mathbf{C}$ by Axiom 6, 2604, and 2605, so $\vdash_{\mathcal{F}^=} \mathbf{D}$.

Suppose $\vdash_{\mathcal{F}^=} \mathbf{D}$. Let \mathcal{T} be the first-order theory whose only constants are the binary predicate constant $=$ and the constants which occur in \mathbf{D} , and whose sole postulate is $\forall x [x = x] \wedge \mathbf{B} \wedge \mathbf{C}$. Let \mathcal{M} be any model for \mathcal{T} . \mathcal{T} is a first-order theory with equality by 2606, so by 2610 \mathcal{T} has an equality-model \mathcal{N} elementarily equivalent to \mathcal{M} . One can easily extend \mathcal{N} to an equality-interpretation \mathcal{N}' for $\mathcal{F}^=$ by arbitrarily interpreting the additional constants of $\mathcal{F}^=$, so $\mathcal{N}' \models \mathbf{D}$ by 2609. Thus $\mathcal{N} \models \mathbf{D}$ since the truth of \mathbf{D} depends only on the constants which occur in it, so $\mathcal{M} \models \mathbf{D}$ since \mathcal{N} and \mathcal{M} are elementarily equivalent. Therefore $\vdash_{\mathcal{T}} \mathbf{D}$ by the Completeness Theorem (2510), so $\vdash_{\mathcal{F}} \forall x [x = x] \wedge \mathbf{B} \wedge \mathbf{C} \supset \mathbf{D}$ by the Deduction Theorem. ■

EXERCISES

X2600. Prove that $\mathcal{F}^=$ is a conservative extension of \mathcal{F} . Assume $=$ is not a predicate symbol of \mathcal{F} .

X2601. Show that the elementary theory of \mathcal{G} of abelian groups which was presented in §21 is a first-order theory with equality.

X2602. Show that if a formulation of \mathcal{F} has two binary predicates $\stackrel{1}{=}$ and $\stackrel{2}{=}$, and is a first-order theory with equality with respect to both of them, then $\vdash \forall x \forall y \bullet x \stackrel{1}{=} y \equiv x \stackrel{2}{=} y$.

X2603. Show that there is no consistent set \mathcal{S} of sentences of $\mathcal{F}^=$ such that all models of \mathcal{S} are equality-models.

Prove the following theorems of $\mathcal{F}^=$:

X2604. Theorem 2607.

X2605. Theorem 2608.

X2606. $\forall x [fx = ffx] \wedge \forall x \forall y [fx = fy \supset x = y] \supset \forall z \exists x \bullet fx = z$

X2607. $\forall x \forall y [fx = fy \supset x = y] \wedge \exists x [\exists u [Pu \wedge fu = x] \wedge \exists v \bullet Qv \wedge fv = x] \supset \exists w \bullet Pw \wedge Qw$

X2608. $\forall x \forall y [x = y] \supset \neg \exists x Px \equiv \forall x Px$

X2609. $\exists x \exists y x \neq y \wedge \exists x \exists y \forall z [z = x \vee z = y] \equiv \exists x \exists y \bullet x \neq y \wedge \forall z \bullet z = x \vee z = y$

Chapter 3

Provability and Refutability

We now turn our attention to the practical problems of proving theorems of first-order logic. If one has unlimited resources and patience, one can simply enumerate proofs until one finds a proof of the wff under consideration (provided, of course, that it actually is a theorem). As a practical matter, however, one usually wishes to find a proof as quickly and easily as possible, so various methods have been devised to facilitate efficient searches for proofs. We shall discuss some of these methods in this chapter, along with various results related to them.

A *refutation* of a wff **B** is a proof that **B** is contradictory; this amounts to a proof of $\sim \mathbf{B}$. Thus, one way to prove a wff **A** is to refute $\sim \mathbf{A}$. This method is proving a wff **A** is often called the indirect method, and it is surprisingly useful, so a number of the procedures we discuss below are refutation procedures rather than proof procedures. Since a refutation of $\sim \mathbf{A}$ provides a proof of **A**, we shall sometimes speak of refutation procedures as proof procedures.

We say that we refute a set \mathcal{S} of sentences when we drive a contradiction from \mathcal{S} . Naturally, such a refutation shows that some finite conjunction of members of \mathcal{S} is contradictory.

§30. Natural Deduction

Proofs of theorems of \mathcal{F} (which is known as a *Hilbert-style* system) are rarely written out in full, since they tend to be long, awkward, and unpleasant to read. In practice, one uses proofs from hypotheses and derived rules of

inference of \mathcal{F} . In this section we introduce a system \mathcal{N} of *natural deduction*,¹ in which it is possible to give fairly natural and well structured proofs, and express the forms of arguments which arise in mathematical practice. Of course, the rules of inference of \mathcal{N} are all derived rules of inference of \mathcal{F} , and this section may be regarded simply as a summary of those derived rules of inference of \mathcal{F} which are most useful for writing our proofs.

Later in this chapter we discuss a variety of methods and logical tools for establishing the validity of wffs of first-order logic. Generally, proofs can be translated from one format into another (though real insight may be required to design algorithms to do this), and once one has found the essential ingredients of a proof, one may wish to express them in a proof which is reasonably comprehensible. The system \mathcal{N} provides one standard for what such a proof should look like.

The exact choice of primitive connectives and quantifiers for this section does not matter very much, but it is natural to take at least \sim , \vee , \wedge , \supset , and f (falsehood) as primitive connectives, and both \forall and \exists as primitive quantifiers. f plays the role of a contradiction in indirect proofs.

Various definitions, such as *free for*, must be adjusted in trivial ways to take account of the fact that \exists is now a primitive quantifier, and we leave this task to the reader. To avoid needless redundancy we shall use M and N as syntactical variables for wffs or for the “null formula” (empty disjunction). When M is null, $M \vee A$ and $A \vee M$ both stand for A . A null formula standing alone may be regarded as an abbreviation for f . In this section we use \mathcal{H} as a syntactical variable for a finite (possibly empty) *sequence* of wffs, and write \mathcal{H}, A for the sequence obtained by appending A to the sequence \mathcal{H} . As in §21, we may say that x is not free in \mathcal{H} when we mean that x is not free in any wff of \mathcal{H} .

A *natural deduction proof* consists of a finite sequence of *lines* of the form $\mathcal{H} \vdash A$. The members of the sequence \mathcal{H} are called *hypotheses*, and the wff A is called the *assertion* of the line. Each line must be inferred from zero or more preceding lines by one of the following rules of inference.

Rules of Inference

Hypothesis Rule (Hyp). Infer $\mathcal{H} \vdash A$ whenever A is a member of \mathcal{H} .

Rule for Expanding or Rearranging Hypotheses. From $\mathcal{H}_1 \vdash A$ infer $\mathcal{H}_2 \vdash A$, provided that every wff in \mathcal{H}_1 is also in \mathcal{H}_2 .

¹Natural deduction was introduced in [Gentzen, 1935]. Also see [Quine, 1950], [Prawitz, 1965], and references cited therein.

Deduction Rule (Ded). From \mathcal{H} , $A \vdash B$ infer $\mathcal{H} \vdash A \supset B$.

Rule P. From $\mathcal{H} \vdash A_1, \dots$, and $\mathcal{H} \vdash A_n$ infer $\mathcal{H} \vdash B$, provided that $[[A_1 \wedge \dots \wedge A_n] \supset B]$ is tautologous.

Negation Rule (Neg.). From $\mathcal{H} \vdash A$, infer $\mathcal{H} \vdash B$, where A is $\sim \forall x C$, $\sim \exists x C$, $\forall x \sim C$, or $\exists x \sim C$, and B is $\exists x \sim C$, $\forall x \sim C$, $\sim \exists x C$, or $\sim \forall x C$, respectively.

Rule of Indirect Proof (IP). From $\mathcal{H}, \sim A \vdash f$ infer $\mathcal{H} \vdash A$.

Rule of Cases (Cases). From $\mathcal{H} \vdash A \vee B$ and $\mathcal{H}, A \vdash C$ and $\mathcal{H}, B \vdash C$ infer $\mathcal{H} \vdash C$.

Rule of Alphabetic Change of Bound Variables (α). From $\mathcal{H} \vdash A$ infer $\mathcal{H} \vdash A'$, where A' is obtained from A upon replacing an occurrence of $\forall x C$ [$\exists x C$] in A by an occurrence of $\forall y S_y^x C$ [$\exists y S_y^x C$], where y is not free in C and y is free for x in C .

Universal Generalization ($\forall G$). From $\mathcal{H} \vdash M \vee A \vee N$ infer $\mathcal{H} \vdash M \vee \forall x A \vee N$, provided that x is not free in \mathcal{H} , M , or N .

Existential Generalization ($\exists G$). Let $A(x)$ be a wff and let t be a term which is free for x in $A(x)$. (t may occur in $A(x)$.) From $\mathcal{H} \vdash M \vee A(t) \vee N$ infer $\mathcal{H} \vdash M \vee \exists x A(x) \vee N$.

Universal Instantiation ($\forall I$). From $\mathcal{H} \vdash \forall x A(x)$ infer $\mathcal{H} \vdash A(t)$, provided that t is a term free for x in $A(x)$.

Rule C. From $\mathcal{H} \vdash \exists x B(x)$ and $\mathcal{H}, B(y) \vdash A$ infer $\mathcal{H} \vdash A$, where y is an individual variable which is free for x in $B(x)$ and which is not free in \mathcal{H} , $\exists x B(x)$ or in A .

Note that the system \mathcal{N} has no axioms. We remark that the Rule for Expanding or Rearranging Hypotheses is often used tacitly and without explicit mention in combination with other rules. Rule $\forall G$ really has four forms:

from $\mathcal{H} \vdash A$ infer $\mathcal{H} \vdash \forall x A$;

from $\mathcal{H} \vdash A \vee C$ infer $\mathcal{H} \vdash \forall x A \vee C$;

from $\mathcal{H} \vdash B \vee A$ infer $\mathcal{H} \vdash B \vee \forall x A$;

from $\mathcal{H} \vdash B \vee A \vee C$ infer $\mathcal{H} \vdash B \vee \forall x A \vee C$.

(In each case, x must not be free in \mathcal{H} , \mathbf{B} , or \mathbf{C} .) Our use of the null formula simply enables us to compress these four statements into one. Rule P can actually be restricted to certain special cases of Rule P (such as certain rules of the system \mathcal{G} of §31), but we shall not discuss that here. Naturally, if one wishes to prove theorems involving $=$, one should add to the rules above certain derived rules of inference of the system discussed in §26. Also, in some contexts it would be natural to include a rule permitting one to infer any previously proved theorem.

The soundness and completeness of \mathcal{N} follow from the corresponding results for \mathcal{F} , though a careful proof must deal with the fact that \exists is a primitive symbol of \mathcal{N} , but not of \mathcal{F} . We leave further consideration of these matters to the reader (Exercises X3003 and X3004).

If $\mathbf{A}(x)$ is a wff in which x occurs free, and one wishes to prove $\exists x \mathbf{A}(x)$, a natural approach is to try to find a term t such that one can prove $\mathbf{A}(t)$, and then derive $\exists x \mathbf{A}(x)$ by $\exists G$. For example, if $\mathbf{A}(x)$ is $[Qy \supset Qx]$, one can prove $\mathbf{A}(y)$ (i.e., $[Qy \supset Qy]$), and from this infer $\exists x \mathbf{A}(x)$ (i.e., $\exists x [Qy \supset Qx]$).

Sometimes this approach will not work, as in the case where $\mathbf{A}(x)$ is $[Q \supset \neg Qa \wedge Qb]$. Nevertheless, in this case one can prove $\mathbf{A}(a) \vee \mathbf{A}(b)$ (i.e., $[Qa \supset \neg Qa \wedge Qb] \vee [Qb \supset \neg Qa \wedge Qb]$, which is tautologous), from which one can infer $\exists x \mathbf{A}(x) \vee \exists x \mathbf{A}(x)$ and hence infer $\exists x \mathbf{A}(x)$ by Rule P. Thus, a natural generalization of the approach to proving $\exists x \mathbf{A}(x)$ mentioned above is to find terms t_1, \dots, t_n such that one can prove $\mathbf{A}(t_1) \vee \dots \vee \mathbf{A}(t_n)$ and from this infer $\exists x \mathbf{A}(x) \vee \dots \vee \exists x \mathbf{A}(x)$, and hence $\exists x \mathbf{A}(x)$.

However, there are cases when even this generalized approach does not quite work. Consider the problem of giving a direct natural deduction proof of $\exists x \forall y \neg Px \supset Py$. One proof is:

$\vdash [Px \supset Py] \vee \neg Py \supset Pz$	Rule P
$\vdash [Px \supset Py] \vee \forall z \neg Py \supset Pz$	$\forall G$
$\vdash [Px \supset Py] \vee \exists x \forall z \neg Px \supset Pz$	$\exists G$
$\vdash \forall y [Px \supset Py] \vee \exists x \forall z \neg Px \supset Pz$	$\forall G$
$\vdash \exists x \forall y [Px \supset Py] \vee \exists x \forall z \neg Px \supset Pz$	$\exists G$
$\vdash \exists x \forall y [Px \supset Py] \vee \exists x \forall y \neg Px \supset Py$	α
$\vdash \exists x \forall y \neg Px \supset Py$	Rule P

Note that in this proof we do not have terms t_1 and t_2 in which y is not free such that we prove $\forall y [Pt_1 \supset Py] \vee \forall y [Pt_2 \supset Py]$. Could there be such terms? Consider an interpretation $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$, where

$$\mathcal{D} = \{a, b\};$$

$\mathcal{J}\mathbf{c} = a$ for all constants \mathbf{c} ;

for each function symbol \mathbf{f}^n

$$(\mathcal{J}\mathbf{f}^n)d_1 \dots d_n = a \text{ for all } d_1, \dots, d_n \text{ in } \mathcal{D};$$

$$(\mathcal{J} P)a = \top, \text{ and } (\mathcal{J} P)b = \perp.$$

Let $\varphi\mathbf{x} = a$ for all individual variables \mathbf{x} ; then $\mathcal{V}_\varphi^M t = a$ for each term t . Thus if y does not occur in a term t , and ψ is the assignment which agrees with φ off y , while $\psi y = b$, then

$$\mathcal{V}_\psi[Pt \supset Py] = \perp, \text{ so } \mathcal{V}_\varphi^M \forall y[Pt \supset Py] = \perp.$$

Thus one cannot prove any disjunction of the form $\bigvee_{i=1}^n \forall y[Pt_i \supset Py]$, where the terms t_i do not contain y .

EXERCISES

Prove the following theorems in \mathcal{N} :

X3000. $\sim \forall \mathbf{x} \mathbf{A} \equiv \exists \mathbf{x} \sim \mathbf{A}$.

X3001. $\sim \exists \mathbf{x} \mathbf{A} \equiv \forall \mathbf{x} \sim \mathbf{A}$.

X3002. $\forall u \forall v \forall w [Puv \vee Pvw] \supset \exists x \forall y Pxy$. (*Hint:* recall the advice given near the end of §21 about proving theorems of the form $\mathbf{A} \supset \mathbf{B}$.)

X3003. Prove that the system \mathcal{N} is sound in the sense of 2303.

X3004. Prove that the system \mathcal{N} is complete in the sense of 2509.

X3005. Add rules of inference for $=$ to \mathcal{N} , and prove that the resulting system is sound and complete in the sense of 2609 and 2612.

X3006. Are any of the rules of inference of \mathcal{N} dependent (non-independent) in the sense of §13?

X3007. Find a system \mathcal{N}' which can be obtained from \mathcal{N} by deleting certain rules of inference, and whose rules of inference are all independent. Prove the independence of these rules.

§31. Gentzen's Theorem

One of the difficulties of analyzing possible proofs of theorems of \mathcal{F} is that the form of a theorem gives few clues concerning what a proof of it might look like. If one seeks to analyze the possible proofs of a theorem B , one must consider the possibility that it was obtained by Modus Ponens from theorems A and $A \supset B$, but it is hard to guess what A might have been. In this section we consider a system \mathcal{G} whose rules of inference are chosen so that one can search in a fairly systematic way for a proof of any specified theorem. The system \mathcal{G} also has the “subformula property”, i.e., every wff which occurs in a proof of a theorem bears some resemblance to some subformula of the theorem. \mathcal{G} is based on ideas introduced by Gentzen [Gentzen, 1935], and is closely related to a simplification of Gentzen's sequent calculus due to Schütte [Schütte, 1950].

The wffs of \mathcal{G} are the same as those of \mathcal{F} . As in §30, we use M and N as syntactical variables for wffs which may be null.

The occurrences of wf parts of a wff which are in the scope of no negations or quantifiers of that wff are called *disjunctive components* of that wff. For example, the disjunctive components of $[[\forall x Fx \vee Gy] \vee [[\sim p \vee \sim Gy] \vee q]]$ are the entire wff, the first occurrence of Gy , and the sole occurrences in the wff of $\forall x Fx$, $[\forall x Fx \vee Gy]$, $\sim p$, $\sim Gy$, $[\sim p \vee \sim Gy]$, q , and $[[\sim p \vee \sim Gy] \vee q]$.

Now we can state the axioms and rules of inference of \mathcal{G} :

Axiom. $\sim A \vee A$, where A is any atomic wff.

Rules of Inference

Disjunction Rules (\vee). Replace a disjunctive component D of a wff by E , where D is $[[A \vee B] \vee C]$ and E is $[A \vee [B \vee C]]$, or D is $[A \vee [B \vee C]]$ and E is $[[A \vee B] \vee C]$, or D is $[A \vee B]$ and E is $[B \vee A]$.

Weakening (Weaken). From B infer $B \vee A$.

Simplification (Simp). From $M \vee [A \vee A]$ infer $M \vee A$.

Negation Introduction (\sim). From $M \vee A$ infer $M \vee \sim \sim A$.

Conjunction Introduction (\wedge). From $M \vee \sim A$ and $M \vee \sim B$ infer $M \vee \sim [A \vee B]$.

Alphabetic Change of Bound Variables (α). (See 2110 for a statement

of this rule.)

Universal Generalization ($\forall G$). From $M \vee A$ infer $M \vee \forall x A$, provided x is not free in M .

Existential Generalization ($\exists G$). From $M \vee \sim S_t^x A$ infer $M \vee \sim \forall x A$, where t is a term free for x in A .

We shall write $\vdash_{\mathcal{G}} A$ to indicate that A is a theorem of \mathcal{G} .

The appropriateness of the names for rules \wedge and $\exists G$ should become clear to the reader who does exercises X3100 and X3101. The \vee rules permit one to move the disjunctive components of a wff around so that the component which is to be modified by a rule of inference is on the right. The need for them could be eliminated by stating the rules of inference in a more general (but less perspicuous) way. We shall often combine applications of the \vee rules with those of other rules, and omit mention of the \vee rules.

EXAMPLE: The following proof in \mathcal{G} illustrates a number of features of the system, including the need for the Simplification rule. The reader may find it enlightening to prove the theorem for himself before looking at the proof below:

(1)	$\vdash \sim Pxy \vee Pxy$	Axiom
(2)	$\vdash \sim Pxy \vee \blacksquare Pxy \vee Pyz$	Weaken
(3)	$\vdash \sim Pyz \vee Pyz$	Axiom
(4)	$\vdash \sim Pyz \vee \blacksquare Pxy \vee Pyz$	Weaken
(5)	$\vdash \sim [Pxy \vee Pyz] \vee \blacksquare Pxy \vee Pyz$	$\wedge : 2, 4$
(6)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \blacksquare Pxy \vee Pyz$	$\exists G$ (thrice)
(7)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \blacksquare Pxy \vee \forall z Pyz$	$\forall G$
(8)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \blacksquare Pxy \vee \exists x \forall z Pxz$	X3101
(9)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \blacksquare \forall y Pxy \vee \exists x \forall z Pxz$	$\forall G$
(10)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \blacksquare \exists x \forall y Pxy \vee \exists x \forall z Pxz$	X3101
(11)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \blacksquare \exists x \forall y Pxy \vee \exists x \forall y Pxy$	α
(12)	$\vdash \sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \exists x \forall y Pxy$	Simp

It can be shown that without loss of the completeness of \mathcal{G} (which we establish below), the Simplification rule can be restricted to the form: from $M \vee \sim \forall x A \vee \sim \forall x A$ to infer $M \vee \sim \forall x A$. Also, one need use it only when the occurrences of $\sim \forall x A$ have been introduced by rule $\exists G$. Therefore the rules Simp and $\exists G$ can be combined into the single rule:

$\exists G$ -Simp. From $M \vee \sim \forall x A \vee \sim S_t^x A$ infer $M \vee \sim \forall x A$, where t is free for x in A .

Note exercise X3102.

3100 Lemma. If $\vdash_{\mathcal{G}} M \vee \sim\sim E$ then $\vdash_{\mathcal{G}} M \vee E$.

Proof: Let P_1, \dots, P_m be a proof in \mathcal{G} of $M \vee \sim\sim E$. We prove by induction on i that if Q_i is the result of replacing zero or more disjunctive components of P_i of the form $\sim\sim D$ by components of the form D , then $\vdash_{\mathcal{G}} Q_i$.

Clearly no axiom contains double negations. Suppose P_i is inferred by some rule, which we shall call rule R , from P_j [and P_k], where $j < i$ [and $k < i$]. If R is not Weakening or \sim , we may apply the inductive hypotheses to P_j [and to P_k] to obtain a wff Q_j [and a wff Q_k] from which Q_i may be inferred by rule R . If R is \sim or Weakening, we proceed in the same way, except that we may omit the application of rule \sim , and we may modify the application of weakening to infer $B \vee D$ instead of $B \vee \sim\sim D$ from B . ■

3101 Theorem. For every wff A of \mathcal{F} , $\vdash_{\mathcal{F}} A$ iff $\vdash_{\mathcal{G}} A$.

Proof: It is easy to see that every theorem of \mathcal{G} is a theorem of \mathcal{F} , since the axioms of \mathcal{G} are all theorems of \mathcal{F} , and the rules of inference of \mathcal{G} are all derived rules of inference of \mathcal{F} .

For the proof in the other direction, we let Γ be the class of all finite sets $\{C_1, \dots, C_n\}$ of sentences such that not $\vdash_{\mathcal{G}} \sim C_1 \vee \dots \vee \sim C_n$, and use Smullyan's Unifying Principle. By virtue of rules \vee , Weaken, and Simp, it does not matter in what order the C_i are listed, and $\{C_1, \dots, C_n\} \cup \{D_1, \dots, D_m\} \in \Gamma$ iff not $\vdash_{\mathcal{G}} \sim C_1 \vee \dots \vee \sim C_n \vee \sim D_1 \vee \dots \vee \sim D_m$ whether or not some C_i and D_j are the same.

Next we verify that Γ is an abstract consistency class. Because \mathcal{G} has the Weakening rule, Γ is clearly closed under subsets.

- (a) If A is atomic, then $\vdash_{\mathcal{G}} M \vee \sim\sim A \vee \sim A$ for any M , so clearly (a) is satisfied.

Conditions (b) and (d)-(f) all have the form: if $S \in \Gamma$ and $D \in S$, then $S \cup \{E_1, \dots, E_n\} \in \Gamma$ (where $1 \leq n \leq 2$). We shall consider the condition in the equivalent (more or less contrapositive) form: assume $D \in S$; if $S \cup \{E_1, \dots, E_n\} \notin \Gamma$, then $S \notin \Gamma$. We may write S as $\{C_1, \dots, C_k, D\}$ (where $k \geq 0$) and $\sim C_1 \vee \dots \vee \sim C_k$ as M . Then in each case the condition takes the form: if $\vdash_{\mathcal{G}} M \vee \sim D \vee \sim E_1 \vee \dots \vee \sim E_n$, then $\vdash_{\mathcal{G}} M \vee \sim D$. Condition (c) can be handled in a similar way. The detailed statements to be verified are the following:

- (b) If $\vdash_{\mathcal{G}} M \vee \sim\sim A \vee \sim A$ then $\vdash_{\mathcal{G}} M \vee \sim\sim A$. Rules \sim and Simp assure this.
- (c) If $\vdash_{\mathcal{G}} M \vee \sim [A \vee B] \vee \sim A$ and $\vdash_{\mathcal{G}} M \vee \sim [A \vee B] \vee \sim B$ then $\vdash_{\mathcal{G}} M \vee \sim [A \vee B]$. Rules \wedge and Simp assure this.
- (d) If $\vdash_{\mathcal{G}} M \vee \sim\sim [A \vee B] \vee \sim\sim A \vee \sim\sim B$, then $\vdash_{\mathcal{G}} M \vee \sim\sim [A \vee B]$. Lemma 3100, \sim , and Simp assure this.
- (e) If $\vdash_{\mathcal{G}} M \vee \sim \forall x A \vee \sim S_t^x A$ for some variable-free term t , then $\vdash_{\mathcal{G}} M \vee \sim \forall x A$. Rules $\exists G$ and Simp assure this.
- (f) If (*) $\vdash_{\mathcal{G}} M \vee \sim\sim \forall x A \vee \sim\sim S_c^x A$ and c is an individual constant which does not occur in M or A , then $\vdash_{\mathcal{G}} M \vee \sim\sim \forall x A$.

To verify (f), note that we can replace c everywhere in the proof of (*) by some new variable y which does not occur in the proof to obtain

$$\begin{array}{ll}
 \vdash_{\mathcal{G}} M \vee \sim\sim \forall x A \vee \sim\sim S_y^x A & \\
 \vdash_{\mathcal{G}} M \vee \forall x A \vee \forall y S_y^x A & 3100 \text{ and } \forall G \\
 \vdash_{\mathcal{G}} M \vee \forall x A \vee \forall x A & \alpha \\
 \vdash_{\mathcal{G}} M \vee \forall x A & \text{Simp} \\
 \vdash_{\mathcal{G}} M \vee \sim\sim \forall x A & \sim
 \end{array}$$

Thus Γ is an abstract consistency class. Now let A be any theorem of \mathcal{F} . If A is not a sentence, let \dot{A} be the result of replacing the free variables $x_1 \dots x_n$ of A by distinct new constants $c_1 \dots c_n$ at all their free occurrences in A . $\vdash_{\mathcal{F}} \dot{A}$ by 2118, so $\{\sim \dot{A}\}$ has no model by 2303 and 2301, so $\{\sim \dot{A}\} \notin \Gamma$ by Smullyan's Unifying Principle (2507), so $\vdash_{\mathcal{G}} \sim\sim \dot{A}$, so $\vdash_{\mathcal{G}} \dot{A}$ by 3100.

Now it remains only to see that a proof of \dot{A} in \mathcal{G} can be transformed into a proof of A . Let y_1, \dots, y_n be distinct variables which do not occur in the given proof of \dot{A} . Replace all occurrences of the x_i in the given proof of \dot{A} by occurrences of the y_i . Then replace all occurrences of the c_i in the resulting proof by occurrences of the x_i . It is easy to see that one thus obtains a proof in \mathcal{G} of a wff which differs from A at most by certain alphabetic changes of bound variables, and by applying Rule α one can then obtain a proof in \mathcal{G} of A . ■

3102 Corollary. If $\vdash_{\mathcal{G}} M \vee A$ and $\vdash_{\mathcal{G}} \sim A \vee N$, then $\vdash_{\mathcal{G}} M \vee N$.

Proof: By 3101 and Rule P (2100) for \mathcal{F} . ■

REMARK. The rule of inference of 3102 is known as the *Cut* rule. Note that it is a generalization of Modus Ponens, since \mathbf{M} may be null. We prefer not to call Cut a derived rule of inference of \mathcal{G} at this stage, since we have not shown how to actually construct a proof of $\mathbf{M} \vee \mathbf{N}$ from given proofs of $\mathbf{M} \vee \mathbf{A}$ and $\sim \mathbf{A} \vee \mathbf{N}$.

It is now easy to describe an alternative approach to 3101 and 3102 which provides a purely syntactic proof of these theorems, and which parallels Gentzen's original argument. Let \mathcal{G}^+ be the result of adding Cut to the primitive rules of inference of \mathcal{G} . It is easy to see that \mathcal{G}^+ and \mathcal{F} have the same theorems, since the axioms of \mathcal{F} are all theorems of \mathcal{G} (Exercise X3104), and the rules of inference of \mathcal{F} are all rules of \mathcal{G}^+ . Then show that by manipulating proofs one can eliminate all applications of Cut from all proofs in \mathcal{G}^+ . This would establish the *Cut-Elimination Theorem* for \mathcal{G}^+ , and show that \mathcal{G} and \mathcal{G}^+ are equivalent systems. Gentzen's *Hauptsatz* (Main Theorem) was the Cut-Elimination Theorem for this system.

\mathcal{G} is called a *Cut-free* system. As we shall mention in §35, Herbrand also discussed a Cut-free system, and part of Herbrand's Theorem is closely related to Gentzen's *Hauptsatz*.

EXERCISES

X3100. Give a direct constructive proof that if $\vdash_{\mathcal{G}} \mathbf{M} \vee \mathbf{A}$ and $\vdash_{\mathcal{G}} \mathbf{M} \vee \mathbf{B}$, then $\vdash_{\mathcal{G}} \mathbf{M} \vee [\mathbf{A} \wedge \mathbf{B}]$, where \wedge as is defined in §12.

X3101. Give a direct constructive proof that if t is free for x in \mathbf{C} and $\vdash_{\mathcal{G}} \mathbf{M} \vee S_t^x \mathbf{C}$, then $\vdash_{\mathcal{G}} \mathbf{M} \vee \exists x \mathbf{C}$, where \exists is defined as in §20.

X3102. Let \mathcal{G}' be the system obtained from \mathcal{G} by deleting rules Simp and $\exists G$, and adding rule $\exists G$ -Simp. Show how to prove $\sim \forall u \forall v \forall w [Puv \vee Pvw] \vee \exists x \forall y Pxy$ in \mathcal{G}' .

X3103. Show that $\vdash_{\mathcal{G}} \sim \mathbf{A} \vee \mathbf{A}$ for every wff \mathbf{A} .

X3104. Show that the axioms of \mathcal{F} are all theorems of \mathcal{G} .

§32. Semantic Tableaux

In this section we introduce a method for refuting sets of sentences which was introduced by Beth [Beth, 1959], Hintikka [Hintikka, 1955], and Schütte [Schütte, 1956]. In Chapters 1 and 2 we dealt with proofs which were finite sequences of wffs, but in this section we shall arrange the occurrences of wffs

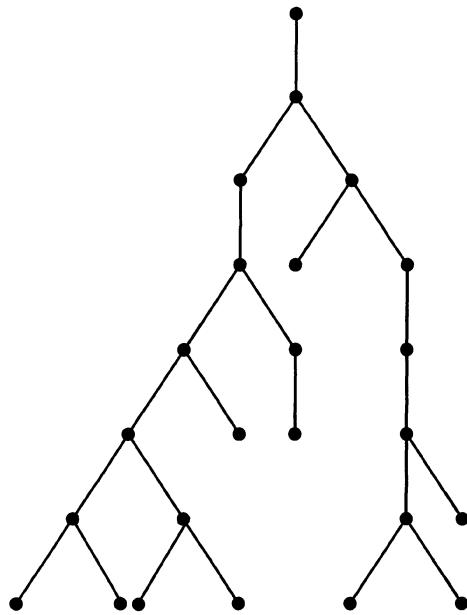


Figure 3.1: A tree

in a refutation as nodes of a tree, and call the resulting structure a *semantic tableau*.

The trees we shall use will be *downward dyadic* trees, meaning that we picture them as having their branches extending downward, and at most two nodes are immediately below any node. An example of such a tree is shown in Figure 3.1, where the nodes are depicted as large dots.

A (*dyadic*) *tree of sentences* is a downward dyadic tree with each node of which is associated some sentence of \mathcal{F} . (We do not use sentences for nodes, since a sentence may have several occurrences in a tree of sentences.) (An occurrence of) a wff \mathbf{A} in such a tree is *below* (an occurrence of) a wff \mathbf{B} iff the node with which \mathbf{A} is associated is below the node with which \mathbf{B} is associated.

DEFINITION. A *semantic tableau* for a finite set \mathcal{S} of sentences is a tree of sentences which can be constructed by the following inductive rules:

- (1) **Initial Hypothesis (Hyp).** A tree of sentences which has just one node, whose sentence is in \mathcal{S} , is a semantic tableau for \mathcal{S} .

- (2) If \mathcal{U} is a semantic tableau for \mathcal{S} , \mathbf{G} is a sentence of \mathcal{U} , and \mathcal{U}' is obtained from \mathcal{U} by appending to a leaf of \mathcal{U} a new node with sentence \mathbf{H} so that \mathbf{H} is below \mathbf{G} in \mathcal{U}' , then \mathcal{U}' is a semantic tableau for \mathcal{S} , provided that \mathbf{H} is obtained from \mathbf{G} by one of the following rules of inference:
- (2a) **Hypothesis (Hyp).** From \mathbf{G} infer \mathbf{H} , where \mathbf{H} is any sentence in \mathcal{S} .
 - (2b) **Double-Negation Elimination ($\sim\sim$).** From $\sim\sim \mathbf{A}$ infer \mathbf{A} .
 - (2c) **Conjunction Elimination to the Left ($\wedge \ell$).** From $\sim [\mathbf{A} \vee \mathbf{B}]$ infer $\sim \mathbf{A}$.
 - (2d) **Conjunction Elimination to the Right ($\wedge r$).** From $\sim [\mathbf{A} \vee \mathbf{B}]$ infer $\sim \mathbf{B}$.
 - (2e) **Universal Instantiation ($\forall I$).** From $\forall x \mathbf{A}$ infer $S_t^x \mathbf{A}$, where t is any variable-free term.
- (3) **Existential Instantiation ($\exists I$).** If \mathcal{U} is a semantic tableau for \mathcal{S} , $\sim \forall x \mathbf{A}$ is a sentence of \mathcal{U} , and \mathcal{U}' is obtained from \mathcal{U} by appending to a leaf of \mathcal{U} a new node with sentence $\sim S_c^x \mathbf{A}$, so that $\sim S_c^x \mathbf{A}$ is below $\sim \forall x \mathbf{A}$ in \mathcal{U}' and c is an individual constant which does not occur in any sentence of \mathcal{U} or \mathcal{S} , then \mathcal{U}' is a semantic tableau for \mathcal{S} .
- (4) **Disjunction Elimination (\vee).** If \mathcal{U} is a semantic tableau for \mathcal{S} , $[\mathbf{A} \vee \mathbf{B}]$ is a sentence of \mathcal{U} with node n , n' is a leaf of \mathcal{U} which is n or is below n , and \mathcal{U}' is obtained from \mathcal{U} by appending below n' two new nodes with sentences \mathbf{A} and \mathbf{B} , respectively, then \mathcal{U}' is a semantic tableau for \mathcal{S} .

A branch in a semantic tableau is *closed* iff it contains a sentence and its negation. We write \square under the terminal leaf of a closed branch. A semantic tableau is *closed* iff all of its branches are closed.

An example of a semantic tableau is displayed in Figure 3.2. Only the sentences associated with the nodes are actually written, with their positions indicating which sentences are below which other sentences. The vertical lines separate nodes which are on different branches of the tree.

$$\exists v \forall x P xv \wedge \forall x [Sx \supset \exists y Qyx] \wedge \forall x \forall y [Pxy \supset \sim Qxy] \supset \exists u \sim Su$$

is proved by constructing a closed semantic tableau for sentences (1)-(4) in Figure 3.2. (The conjunction of (1)-(4) is equivalent to the negation of the sentence to be proved.)

(1)	$\exists v \forall x P xv$	hyp
(2)	$\forall x [Sx \supset \exists y Q yx]$	hyp
(3)	$\forall x \forall y [Pxy \supset \sim Qxy]$	hyp
(4)	$\forall u Su$	hyp
(5)	$\sim \sim \forall x Pxa$	$\exists I : 1$
(6)	$\forall x Pxa$	$\sim \sim : 5$
(7)	$\sim Sa \vee \exists y Qya$	$\forall I : 2$
(8)	$\sim Sa \vee : 7$	$\vee : 7$
(10)	$Sa \quad \forall I : 4$	$\exists I : 9$
	$\square \quad 8, 10$	$\sim \sim : 11$
	(9) $\exists y Qya$	$\vee : 7$
	(11) $\sim \sim Qba$	$\exists I : 9$
	(12) Qba	$\sim \sim : 11$
	(13) $\forall y [Pby \supset \sim Qby]$	$\forall I : 3$
	(14) $\sim Pba \vee \sim Qba$	$\forall I : 13$
	(15) $\sim Pba \vee : 14$	$\vee : 14$
	(16) $\sim Qba \vee : 14$	$\square \quad 12, 16$
	(17) $Pba \quad \forall I : 6$	
	$\square \quad 15, 17$	

Figure 3.2: A semantic tableau

One can speed up the construction of semantic tableaux by applying several rules of inference as one step, and this process can be systematized by formulating, justifying, and applying derived rules of inference. For example, the inference from (1) to (6) in Figure 3.2 can be justified by a slight generalization of rule $\exists I$. Similarly, $\wedge l$ and $\wedge r$ can be generalized to permit inferring **A** and **B** from $[A \wedge B]$ (which is defined as in §10, of course). These generalized rules are used in the abbreviated tableau in Figure 3.3. There is a sense in which modus ponens is a derived rule of inference, since a branch containing **A** and $[\sim A \vee B]$ can be split into a closed branch containing $\sim A$, and another branch containing **B**, as illustrated in Figure 3.4. We leave it to the reader to pursue this subject further (Exercise X3200).

3200 Theorem. Let \mathcal{S} be a set of sentences of \mathcal{F} . \mathcal{S} has no model iff \mathcal{S} has a closed semantic tableau.

Proof: It is easy to see that if \mathcal{S} is a set of sentences which has a model and \mathcal{U} is a semantic tableau for \mathcal{S} , then \mathcal{U} has a branch \mathcal{B} such that the set of sentences associated with (nodes of) \mathcal{B} has a model. This is seen by induction on the construction of \mathcal{U} , considering each of the rules (1)-(4) in turn, and we leave the details to the reader. Clearly, if \mathcal{B} is a branch of

(1)	$\forall u \forall v \forall w [Puv \vee Pvw] \wedge \forall x \exists y \sim Pxy$	<i>hyp</i>
(2)	$\forall u \forall v \forall w [Puv \vee Pvw]$	$\wedge 1 : 1$
(3)	$\forall x \exists y \sim Pxy$	$\wedge r : 1$
(4)	$\exists y \sim Pay$	$\forall : 3$
(5)	$\sim Pab$	$\exists E : 4$
(6)	$\exists y \sim Pby$	$\forall A : 3$
(7)	$\sim Pbc$	$\exists E : 6$
(8)	$Pab \vee Pbc$	$\forall I (thrice) : 2$
(9)	$Pab \vee : 8$	
	$\square \quad 5, 9$	(10) $Pbc \vee : 8$
		$\square \quad 7, 10$

Figure 3.3: An abbreviated tableau

\mathcal{U} which has a model (in the sense above), then \mathcal{B} cannot be closed, so \mathcal{U} cannot be closed. Consequently, if \mathcal{S} has a closed tableau, then \mathcal{S} can have no model.

To establish that every set of sentences \mathcal{S} which has no model has a closed tableau, we use Smullyan's Unifying Principle (2507). Let $\Gamma = \{ \mathcal{S} \mid \mathcal{S} \text{ is a set of sentences which has no closed tableau} \}$.

Clearly, if a finite set \mathcal{S} of sentences has a closed tableau, every finite superset of \mathcal{S} has a closed tableau. From this it easily follows that Γ is closed under subsets.

It is straightforward to establish that Γ is an abstract consistency class by checking conditions (a)-(f) of the definition in §25. Note that when \mathcal{G} is a set of sentences, $\mathcal{G} \notin \Gamma$ means that \mathcal{G} has a closed tableau. Thus to check condition (e) (in contrapositive form), it suffices to show that if, for some variable-free term t , $\mathcal{S} \cup \{\mathcal{S}_t^x A\}$ has a closed tableau \mathcal{Q} , and $\forall x A \in \mathcal{S}$, then \mathcal{S} also has a closed tableau \mathcal{Q}' . Here \mathcal{Q}' can be constructed from \mathcal{Q} by placing a new node, with sentence $\forall x A$, above the origin of \mathcal{Q} , so that the

		⋮		
	(6) A			
	(7) ~A ∨ B			
(8) ~A	∨ : 7		(9) B	∨ : 7
□	6, 8			

Figure 3.4: Modus ponens in a tableau

new node becomes the origin of \mathcal{Q}' . We leave it to the reader to check the remaining conditions.

Now let \mathcal{S} be a set of sentences which has no model. By the Compactness Theorem (2514), \mathcal{S} has a finite subset \mathcal{S}_0 which has no model. By adjusting the language, if necessary, we may assume that \mathcal{S}_0 is a set of sentences in a countable language with infinitely many individual constants. Therefore, by Smullyan's Unifying Principle, $\mathcal{S}_0 \notin \Gamma$, so \mathcal{S}_0 , and hence \mathcal{S} , has a closed tableau. ■

Thus the method of semantic tableaux provides a sound and complete refutation procedure for first-order logic. While it is in some ways a simple-minded and inefficient procedure, constructing a closed tableau for a sentence $\sim A$ often seems easier than directly finding a proof for A in the system \mathcal{F} of Chapter 2.

Indeed, a semantic tableau can be constructed in a very systematic way. For this purpose let us combine $\wedge l$ and $\wedge r$ into the single rule \wedge : From $\sim [A \vee B]$ infer $\sim A$ and $\sim B$. Now let C be any sentence in a semantic tableau. If C is a literal, no rule of inference can be applied to C . If C is a non-literal, exactly one rule of inference can be applied to C , and unless C has the form $\forall x A$, there is no point in applying that rule to C more than once. The rule $\forall I$ may have to be applied more than once to a wff of the form $\forall x A$, so that x is instantiated with a number of different terms, as in Figure 3.3. However, the choice of these terms can be restricted as described in Exercise X3203. When function symbols are not present, this restricts the choice of terms to the finite set of individual constants which have already occurred on the current branch of the tableau (unless that set is empty). Of course, the set of available constants keeps expanding with applications of $\exists I$.

EXERCISES

X3200. Formulate, justify, and illustrate some derived rules of inference for semantic tableaux.

X3201. Use the method of semantic tableau to establish the validity of the wffs in exercises X2106- X2113.

X3202. Show by a purely syntactic argument (as opposed to the semantic one given in the proof of Theorem 3200) that if a set \mathcal{S} of sentences has a closed semantic tableau, then \mathcal{S} is inconsistent. Indeed, give an explicit method for transforming a closed tableau for \mathcal{S} into a refutation of \mathcal{S} in the

system of §21.

X3203. Show that if a set of sentences has a closed semantic tableau, then it has one satisfying the following condition: Let $S_t^x A$ be any wff inferred from a wff of the form $\forall x A$ in the tableau, and let B be the set of sentences which are above $S_t^x A$ in the tableau. If at least one individual constant occurs in some sentence of B , then every individual and function constant which occurs in t also occurs in some sentence of B .

§33. Skolemization

It will be recalled that when a wff is put into prenex normal form, its positive (universal) quantifiers come out as universal quantifiers, and its negative (universal) quantifiers come out as existential quantifiers. Thus, when one speaks of wffs of \mathcal{F} (without abbreviations), so that only universal quantifiers occur, one may regard positively occurring quantifiers as essentially universal, and negatively occurring quantifiers as essentially existential.

DEFINITION. A wff A of \mathcal{F} is *universal* iff every occurrence of a quantifier in A is positive in A .

We now wish to associate with each sentence B of a countable formulation of \mathcal{F} a universal sentence $\star B$ of an expansion \mathcal{F}^* of \mathcal{F} . We let \mathcal{F}^* be the result of adding to \mathcal{F} a denumerable sequence $\bar{g}_1^n, \bar{g}_2^n, \dots, \bar{g}_j^n, \dots$ of new n -ary function constants for each natural number $n \geq 0$. (When $n = 0$, \bar{g}_j^n is an individual constant.) We may refer to these constants as *existential constants*. We fix an enumeration $B_1, B_2, \dots, B_k, \dots$ of the sentences of \mathcal{F} , and define $\star B_1, \star B_2, \dots, \star B_k, \dots$ in order. When, in the definition of $\star B_k$, we refer to the first “new n -ary existential constant”, we mean the first of the constants \bar{g}_j^n which has not been used in the definition of $\star B_1, \dots, \star B_{k-1}$, or used previously in the definition of $\star B_k$.

To define $\star B$ for a given sentence B of \mathcal{F} , first let $\star^0 B$ be a rectified sentence (see §22) obtained from B by deleting vacuous quantifiers and making alphabetic changes of bound variables; thus $\vdash B \equiv \star^0 B$. Let m be the number of negative occurrences of quantifiers in $\star^0 B$. We inductively define $\star^i B$, for $0 < i \leq m$, and let $\star B$ be $\star^m B$. Given $\star^i B$, let $\forall y D$ be the first (in left-to-right order) negative occurrence of a quantifier with its scope in $\star^i B$, and let z_1, \dots, z_n (where $n \geq 0$) be the free variables of $\forall y D$. Let \bar{g}^n be the first new n -ary existential constant, and let $\star^{i+1} B$ be the result of replacing $\forall y D$ by $S_{\bar{g}^n z_1 \dots z_n}^y D$ in $\star^i B$. (Naturally, if $n = 0$, $\forall y D$

is replaced by $S_{\bar{g}^n z_1 \dots z_n}^x D$.)

This completes the definition of $\star B$. Clearly $\star^i B$ contains $m - i$ negative occurrences of quantifiers, so $\star B$ is a universal sentence.

We shall refer to $\star B$ as the *Skolemized form*², or the *functional form (for satisfiability)* of B . Since Skolem introduced [Skolem, 1928], and Herbrand extensively exploited [Herbrand, 1930], the idea of replacing existential quantifiers by function symbols, the existential constants are sometimes referred to *Skolem functors* or *Herbrand functors*.

EXAMPLE 1: If B is $\forall z \exists y Pzy$, $\star B$ is $\forall z \sim \sim Pz\bar{g}^1 z$.

EXAMPLE 2: (We shall introduce a^0 , g^1 , and h^1 as existential constants. For greater legibility we shall write a^0 as a , $g^1 z$ as (gz) , and $h^1 z$ as (hz) .)

$$\begin{aligned} B &= \exists x \forall z [Pzx \equiv \forall u \forall v \cdot Pzu \supset \cdot Puv \supset \sim Pvz] \\ \star^0 B &= \exists x \forall z [[Pzx \supset \forall u \forall v \cdot Pzu \supset \cdot Puv \supset \sim Pvz] \\ &\quad \wedge \cdot \forall y \forall w [Pzy \supset \cdot Pyw \supset \sim Pwz] \supset Pzx] \\ \star^1 B &= \sim \sim \forall z [[Pza \supset \forall u \forall v \cdot Pzu \supset \cdot Puv \supset \sim Pvz] \\ &\quad \wedge \cdot \forall y \forall w [Pzy \supset \cdot Pyw \supset \sim Pwz] \supset Pza] \\ \star^2 B &= \sim \sim \forall z [[Pza \supset \forall u \forall v \cdot Pzu \supset \cdot Puv \supset \sim Pvz] \\ &\quad \wedge \cdot \forall w [Pz(gz) \supset \cdot P(gz)w \supset \sim Pwz] \supset Pza] \\ \star B &= \sim \sim \forall z [[Pza \supset \forall u \forall v \cdot Pzu \supset \cdot Puv \supset \sim Pvz] \\ &\quad \wedge \cdot [Pz(gz) \supset \cdot P(gz)(hz) \supset \sim P(hz)z] \supset Pza] \end{aligned}$$

It should be observed that if the negative quantifier $\forall w$ had been eliminated before $\forall y$, then w would have been replaced first by $(h^2 zy)$, and then by $(h^2 z(g^1 z))$ when $\forall y$ was eliminated. The simplicity of $(h^1 z)$ as compared with $(h^2 z(g^1 z))$ is the chief reason for eliminating the negative quantifiers in left-to-right order.

3300 Lemma. For any sentence B of \mathcal{F} , $\vdash_{\mathcal{F}^*} (\star B) \supset B$.

Proof: In the definition of $\star^{i+1} B$, $\vdash \forall y D \supset S_{\bar{g}^n z_1 \dots z_n}^x D$, since the fact that $\star^i B$ is rectified assures that $\bar{g}^n z_1 \dots z_n$ is free for y in D . Hence $\vdash \star^{i+1} B \supset \star^i B$ for $0 \leq i < m$ by the Substitutivity of Implication (2105), so $\vdash (\star B) \supset B$ by Rule P. ■

²This is to be distinguished from the *Skolem normal form* of a wff, in which the sentence is in prenex normal form, has at least one existential quantifier, and all existential quantifiers precede all universal quantifiers; see [Church, 1956, pp. 224-227]

3301 Theorem. Let \mathcal{G} be a set of sentences of \mathcal{F} and let $\mathcal{G}^* = \{\star\mathbf{B} \mid \mathbf{B} \in \mathcal{G}\}$. Then \mathcal{G}^* has a model iff \mathcal{G} has a model, and $\mathcal{F}^* \cup \mathcal{G}^*$ is a conservative extension $\mathcal{F} \cup \mathcal{G}$.

EXAMPLE: Let \mathcal{G} be the set $\{G1, \dots, G7\}$ of postulates for the elementary theory of additive abelian groups discussed near the beginning of §21. Here \mathcal{G}^* is the result of replacing $G6$ by $\star G6$, which is (modulo deletion of a double negation) $\forall x \cdot x + gx = 0$. Traditionally one writes gx as $-x$, and abbreviates $\forall x \cdot x + (-x) = 0$ as $\forall x \cdot x - x = 0$.

We could have used the single postulate $\exists z \cdot \forall x[x + z = x] \wedge \forall x \exists y[x + y = z]$ in place of $G5$ and $G6$ in \mathcal{G} . The Skolemized form of this sentence could be written as $\forall x[x + 0 = x] \wedge \forall x[x - x = 0]$.

REMARK. Since $\mathcal{F}^* \cup \mathcal{G}^*$ is a conservative extension of $\mathcal{F} \cup \mathcal{G}$, one knows that one can use the convenient notation introduced by the Skolem functions without changing the mathematical content of the theory.

Proof: By Lemma 3300 and the Soundness Theorem (2303), every model of \mathcal{G}^* is a model of \mathcal{G} , and $\mathcal{F}^* \cup \mathcal{G}^*$ is an extension of $\mathcal{F} \cup \mathcal{G}$.

If \mathcal{G} has a model, it has a denumerable model by the Löwenheim-Skolem Theorem (2513), so let $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$ be any denumerable model of the first order theory $\mathcal{F} \cup \mathcal{G}$. We shall extend \mathcal{M} to a model $\mathcal{M}^* = \langle \mathcal{D}, \mathcal{J}^* \rangle$ of $\mathcal{F}^* \cup \mathcal{G}^*$ with the same domain of individuals such that for any sentence \mathbf{B} of \mathcal{F} , $\mathcal{M} \models \mathbf{B}$ iff $\mathcal{M}^* \models \star\mathbf{B}$. Since the domain of individuals of \mathcal{M} is in one-one correspondence with the natural numbers, we may assume it is well-ordered. (This will enable us to avoid any appeal to the Axiom of Choice in the argument below.) Let 0 denote the least individual of \mathcal{D} under this ordering.

For each existential constant \bar{g}^n of \mathcal{F}^* there is a unique sentence \mathbf{B} of \mathcal{F} such that \bar{g}^n occurs in $\star\mathbf{B}$. We extend \mathcal{J} to \mathcal{J}^* in an inductive fashion by considering the existential constants in the order in which they were used in defining the sequence $\star\mathbf{B}_1, \star\mathbf{B}_2, \dots, \star\mathbf{B}_k, \dots$, where \mathbf{B}_k is the k th sentence of \mathcal{F} . As we extend \mathcal{J} to \mathcal{J}^* we show that if \mathbf{B} is any sentence of \mathcal{F} and $\star^i\mathbf{B}$ is defined and \mathcal{J} has been extended to an interpretation function \mathcal{J}' whose domain contains all existential constants of $\star^i\mathbf{B}$, then $\mathcal{V}^{\langle \mathcal{D}, \mathcal{J}' \rangle} \mathbf{B} = \mathcal{V}^{\langle \mathcal{D}, \mathcal{J} \rangle} \mathbf{B}$.

Suppose $\star^{i+1}\mathbf{B}$ was obtained by replacing a wf part $\forall y \mathbf{D}$ of $\star^i\mathbf{B}$ by $\underline{\mathbf{S}}_{\bar{g}^n \mathbf{z}_1 \dots \mathbf{z}_n}^y \mathbf{D}$, where $\mathbf{z}_1, \dots, \mathbf{z}_n$ are the free variables of $\forall y \mathbf{D}$. We may suppose that \mathcal{J} has already been extended to \mathcal{J}' , which is already defined on all constants of $\star^i\mathbf{B}$, so $\mathcal{V}^{\langle \mathcal{D}, \mathcal{J}' \rangle} \star^i\mathbf{B} = \mathcal{V}^{\langle \mathcal{D}, \mathcal{J} \rangle} \mathbf{B}$, but $\mathcal{J}' \bar{g}^n$ is

not defined. We extend \mathcal{J}' to a function \mathcal{J}'' which also provides an interpretation for $\bar{\mathbf{g}}^n$. Let d_1, \dots, d_n be arbitrary elements of \mathcal{D} ; we must define $(\mathcal{J}''\bar{\mathbf{g}}^n)(d_1, \dots, d_n)$. Let φ be an assignment such that $\varphi z_j = d_j$ for $1 \leq j \leq n$. Let $(\mathcal{J}''\bar{\mathbf{g}}^n)(d_1, \dots, d_n)$ be the least member d of \mathcal{D} such that $\mathcal{V}_\psi^{\langle \mathcal{D}, \mathcal{J}' \rangle} \mathbf{D} = F$, where ψ agrees with φ off \mathbf{y} and $\psi \mathbf{y} = d$, if such an element exists; if not, let $d = 0$. This completes the definition of $\mathcal{J}''\bar{\mathbf{g}}^n$.

We next show that for any assignment φ ,

$$\mathcal{V}_\varphi^{\langle \mathcal{D}, \mathcal{J}'' \rangle} \forall \mathbf{y} \ \mathbf{D} = \mathcal{V}_\varphi^{\langle \mathcal{D}, \mathcal{J}'' \rangle} S_{\bar{\mathbf{g}}^n z_1 \dots z_n}^{\mathbf{y}} \mathbf{D}.$$

If $\mathcal{V}_\varphi S_{\bar{\mathbf{g}}^n z_1 \dots z_n}^{\mathbf{y}} \mathbf{D} = F$, then $\mathcal{V}_\varphi \forall \mathbf{y} \ \mathbf{D} = F$, since $\vdash \forall \mathbf{y} \ \mathbf{D} \supset S_{\bar{\mathbf{g}}^n z_1 \dots z_n}^{\mathbf{y}} \mathbf{D}$, and every theorem is valid. Suppose $\mathcal{V}_\varphi \forall \mathbf{y} \ \mathbf{D} = F$. Then there is an assignment ψ which agrees with φ off \mathbf{y} such that $\mathcal{V}_\psi \mathbf{D} = F$; let ψ be such an assignment with the least possible value for $\psi \mathbf{y}$. Then $\mathcal{V}_\varphi \bar{\mathbf{g}}^n z_1 \dots z_n = (\mathcal{J}''\bar{\mathbf{g}}^n)(\varphi z_1, \dots, \varphi z_n) = \psi \mathbf{y}$ by the definition of $\mathcal{J}''\bar{\mathbf{g}}^n$ (since $\mathcal{V}_\psi^{\langle \mathcal{D}, \mathcal{J}' \rangle} \mathbf{D} = \mathcal{V}_\psi^{\langle \mathcal{D}, \mathcal{J}'' \rangle} \mathbf{D}$), so $\mathcal{V}_\varphi S_{\bar{\mathbf{g}}^n z_1 \dots z_n}^{\mathbf{y}} \mathbf{D} = \mathcal{V}_\psi \mathbf{D} = F$ by the Substitution-Value Theorem (2302).

It is now clear that $\langle \mathcal{D}, \mathcal{J}'' \rangle \models \forall z_1 \dots \forall z_n \bullet \forall \mathbf{y} \ \mathbf{D} \equiv S_{\bar{\mathbf{g}}^n z_1 \dots z_n}^{\mathbf{y}} \mathbf{D}$, but by the Substitutivity of Equivalence (2106), $\vdash \forall z_1 \dots \forall z_n [\forall \mathbf{y} \ \mathbf{D} \equiv S_{\bar{\mathbf{g}}^n z_1 \dots z_n}^{\mathbf{y}} \mathbf{D}] \supset \star^i \mathbf{B} \equiv \star^{i+1} \mathbf{B}$. Since every theorem is valid, $\langle \mathcal{D}, \mathcal{J}'' \rangle \models \star^i \mathbf{B} \equiv \star^{i+1} \mathbf{B}$. Since $\bar{\mathbf{g}}^n$ does not occur in $\star^i \mathbf{B}$, it is easy to see that $\mathcal{V}^{\langle \mathcal{D}, \mathcal{J}'' \rangle} \star^{i+1} \mathbf{B} = \mathcal{V}^{\langle \mathcal{D}, \mathcal{J}'' \rangle} \star^i \mathbf{B} = \mathcal{V}^{\langle \mathcal{D}, \mathcal{J}' \rangle} \star^i \mathbf{B} = \mathcal{V}^{\langle \mathcal{D}, \mathcal{J} \rangle} \mathbf{B}$, and the induction is complete.

Naturally when $\star^i \mathbf{B}$ is $\star \mathbf{B}$, no further extension of \mathcal{J}' to \mathcal{J}^* affects the value of \mathbf{B} , so $\mathcal{V}^{\mathcal{M}^*} \star \mathbf{B} = \mathcal{V}^{\langle \mathcal{D}, \mathcal{J}' \rangle} \star \mathbf{B} = \mathcal{V}^{\mathcal{M}} \mathbf{B}$. Consequently, for any denumerable model \mathcal{M} of $\mathcal{F} \cup \mathcal{G}$, \mathcal{M}^* is a model of $\mathcal{F}^* \cup \mathcal{G}^*$.

To complete the proof of the theorem, suppose \mathbf{C} is any wff of \mathcal{F} such that $\vdash_{\mathcal{F}^* \cup \mathcal{G}^*} \mathbf{C}$. Let \mathcal{M} be any denumerable model of $\mathcal{F} \cup \mathcal{G}$. Then \mathcal{M}^* is a model of $\mathcal{F}^* \cup \mathcal{G}^*$, so $\mathcal{M}^* \models \mathbf{C}$ by the Soundness Theorem (2303b), so $\mathcal{M} \models \mathbf{C}$. Thus \mathbf{C} is valid in all denumerable models of $\mathcal{F} \cup \mathcal{G}$, so $\vdash_{\mathcal{F} \cup \mathcal{G}} \mathbf{C}$ by the Generalized Completeness Theorem (2509). ■

3302 Corollary. A sentence \mathbf{B} of \mathcal{F} is satisfiable iff $\star \mathbf{B}$ is satisfiable.

Proof: Let $\mathcal{G} = \{\mathbf{B}\}$.

EXAMPLE: Let $B = \forall x[\exists y Pxy \supset \exists y Pxy]$, where P is a binary predicate constant. B is valid, hence satisfiable. Hence the sentence $\forall x[\exists y Pxy \supset Px(gx)]$ is also satisfiable.

3303 Corollary. Let \mathbf{A} be any wff, and let $\overline{\mathbf{A}}$ be its universal closure. \mathbf{A} is valid iff $\star \sim \overline{\mathbf{A}}$ is unsatisfiable.

Proof: The following conditions are equivalent:

- (a) \mathbf{A} is valid.
- (b) $\overline{\mathbf{A}}$ is valid. (by 2301 c)
- (c) $\sim \overline{\mathbf{A}}$ is unsatisfiable. (2301 a)
- (d) $\star \sim \overline{\mathbf{A}}$ is unsatisfiable. (by 3302) ■

In addition to the method of Skolemization \star which is discussed above, there are two other methods of performing Skolemization which have been discussed in the literature. We shall call them \mathcal{S}_1 and \mathcal{S}_2 . In each method one replaces wf parts $\exists y M(y)$ of a rectified sentence C by $M(gx^1 \dots x^n)$ where $n \geq 0$, g is a new function (or individual) constant, and the arguments x^1, \dots, x^n of g are certain variables. In method \mathcal{S}_1 , the arguments of g are the variables x^i such that there is an occurrence of $\forall x^i$ in C with $\exists y M$ in its scope. In method \mathcal{S}_2 , the arguments of g are the variables x^i such that there is an occurrence of $\forall x^i$ in C with the minimal scope of $\exists y M$ in its minimal scope. Of course, in method \star , the arguments of g are the free variables of $\exists y M$. For each method, the result of eliminating all existential quantifiers from C in this way is called the Skolemization $\mathcal{S}(C)$ of C , and it can be shown that C has a model if and only if $\mathcal{S}(C)$ has a model. We shall write $\mathcal{S}(C)$ as $\mathcal{S}_1(C)$, $\mathcal{S}_2(C)$, or $\star(C)$ when we need to specify which Skolemization method is used. To see that the three methods can give different results, let C be the sentence

$$\forall x \exists w \forall u [\exists z Pwz \wedge [Mxuw \vee \exists y Quy]].$$

$$\mathcal{S}_1(C) \text{ is } \forall x \forall u [P(fx)(gxu) \wedge [Mxu(fx) \vee Qu(hxu)]].$$

$$\mathcal{S}_2(C) \text{ is } \forall x \forall u [P(fx)(gx) \wedge [Mxu(fx) \vee Qu(hxu)]].$$

$$\star(C) \text{ is } \forall x \forall u [P(fx)(gx) \wedge [Mxu(fx) \vee Qu(hu)]].$$

The most commonly used method of Skolemization is \mathcal{S}_1 , which was introduced by Skolem [Skolem, 1928] (see pp. 517-518 of [van Heijenoort, 1967], though he applied it only to wffs in prenex normal form. In [Herbrand, 1930], Herbrand used both methods \mathcal{S}_1 and \mathcal{S}_2 (see pp. 138, 151, and 154-154 of [Herbrand, 1971] or pp. 533 and 542-544 of [van Heijenoort, 1967]), and applied them to wffs which need not be in prenex normal form.

Herbrand used \mathcal{S}_1 in defining Property C of wffs, and \mathcal{S}_2 in defining Property B. Method \star is used in [Andrews, 1981]. When applied to formulas in antiprenex (miniscope) form, \mathcal{S}_2 seems to represent a compromise between \mathcal{S}_1 and \star . As illustrated above, in some cases \star gives the Skolem functions fewer arguments than \mathcal{S}_1 or \mathcal{S}_2 , since x^i may not occur free in $\exists y M$ even though $\exists y M$ is in the scope of $\forall x^i$. Therefore $\star(C)$ may be simpler than $\mathcal{S}_1(C)$ or $\mathcal{S}_2(C)$.

§34. Refutations of Universal Sentences

In this section we extend the Ground Resolution Theorem (1600) to first-order logic. The proof procedure whose completeness is established by Theorem 3401 below has many of the elements of the Resolution method for automated theorem proving introduced in [Robinson, 1965]. However, the Resolution method has an additional very important element, the use of a unification algorithm to generate the substitutions which are needed to prepare for applications of the Simplification and Cut rules discussed below. Such an algorithm will be discussed in §36.

When one uses the semantic tableau method (see §32), one can introduce many unnecessary nodes by performing universal instantiations which are not actually needed for the proof. Of course, when one has a universally quantified formula to instantiate, one may not yet have enough information to decide intelligently which instantiation terms to use. One approach to this problem is to simply instantiate universal quantifiers with variables, and make appropriate substitutions for these variables later. However, if one is not careful how one handles existential instantiations in this context, one can destroy the soundness of the refutation procedure. A way to avoid this problem is to Skolemize the sentence to be refuted before starting the refutation. When this is done, one simply needs to refute a universal sentence, and existential instantiations never need be made.

Another feature of the semantic tableau method which is sometimes disadvantageous is the division into cases (two branches of the tree) whenever one encounters a disjunction. When one encounters many disjunctions, and does not know which wffs are actually needed for the proof, this procedure can lead to a very redundant tree with a large number of branches. An alternative to this method of handling disjunctions is to use the Cut rule (see §31).

In this section we discuss a method for refuting sets of universal sentences which incorporates these ideas. We also consider extending the pre-processing so as to eliminate conjunctions as well as essentially existential quantifiers.

For this section, it is convenient to use a formulation of first-order logic which we call \mathcal{F}_f , whose primitive symbols are those of \mathcal{F} plus a 0-ary propositional connective f which is false in all interpretations. (With appropriate minor modifications, the theorem below also applies to formulations of \mathcal{F} which do not contain f .) It is readily verified that Smullyan's Unifying Principle (2507) applies to \mathcal{F}_f if one modifies condition (a) in the definition of abstract consistency class to read:

(a') $f \notin \mathcal{S}$, and if \mathbf{A} is atomic, then $\mathbf{A} \notin \mathcal{S}$ or $\sim \mathbf{A} \notin \mathcal{S}$.

As in §31, occurrences of wf parts of a wff \mathbf{A} of \mathcal{F}_f which are in the scope of no negations or quantifiers of \mathbf{A} are called disjunctive components of \mathbf{A} . We shall use \square as a notation for the empty disjunction or "null formula". (Technically, this is a slightly different usage of \square from that in §16, though the two usages are closely related; they simply occur in slightly different contexts.) As in §30, we use M and N as syntactical variables which may take as values \square as well as wffs, and we regard $\mathbf{A} \vee \square$ and $\square \vee \mathbf{A}$ as abbreviations for \mathbf{A} . (For example, we may regard $\sim\sim \mathbf{B}$ as having the form $M \vee \sim\sim \mathbf{B}$.) \square standing alone may be regarded as an abbreviation for f .

DEFINITION. An \mathcal{R} -derivation of \mathbf{G} from a set \mathcal{S} of universal sentences is a finite sequence $\mathbf{B}_1, \dots, \mathbf{B}_k$ of wffs such that $\mathbf{B}_k = \mathbf{G}$, and for each i ($1 \leq i \leq k$), $\mathbf{B}_i \in \mathcal{S}$ or \mathbf{B}_i is inferred from preceding members of the sequence by one of the following rules of inference:

(R1) Disjunction Rules. To replace a disjunctive component \mathbf{D} of a wff by \mathbf{E} , where

- \mathbf{D} is $[[\mathbf{A} \vee \mathbf{B}] \vee \mathbf{C}]$ and \mathbf{E} is $[\mathbf{A} \vee [\mathbf{B} \vee \mathbf{C}]]$, or
- \mathbf{D} is $[\mathbf{A} \vee [\mathbf{B} \vee \mathbf{C}]]$ and \mathbf{E} is $[[\mathbf{A} \vee \mathbf{B}] \vee \mathbf{C}]$, or
- \mathbf{D} is $[\mathbf{A} \vee \mathbf{B}]$ and \mathbf{E} is $[\mathbf{B} \vee \mathbf{A}]$.

(R2) Simplification. From $\mathbf{M} \vee \mathbf{A} \vee \mathbf{A}$ to infer $\mathbf{M} \vee \mathbf{A}$.

(R3) Substitution. From \mathbf{A} to infer $S_{t_1 \dots t_n}^{x_1 \dots x_n} \mathbf{A}$, where x_1, \dots, x_n are distinct individual variables and t_1, \dots, t_n are terms such that t_i is free for x_i in \mathbf{A} for $1 \leq i \leq n$.

(R4) **Cut.** From $M \vee A$ and $N \vee \sim A$ to infer $M \vee N$.

(R5) **Negation Elimination.** From $M \vee \sim\sim A$ to infer $M \vee A$.

(R6) **Conjunction Elimination.** From $M \vee \sim [A \vee B]$ to infer $M \vee \sim A$ and $M \vee \sim B$.

(R7) **Universal Instantiation.** From $M \vee \forall x A$ to infer $M \vee A$.

REMARK. It might be supposed that the α Rule is also needed for a complete refutation system, to assure that x is distinct from the free variables of M in R7 so that after applying R7 one can substitute for the free occurrences of x in A without involving those in M . However, one can use R3 to substitute new variables for the free variables of $M \vee \forall x A$ before applying R7.

DEFINITION. We shall write $S \vdash_R E$ to indicate that there is an R -derivation of E from S . An R -derivation of \square from S will be called an R -refutation of S . If $S \vdash_R \square$ we shall say that S is R -refutable.

3400 Theorem. Let S be any set of universal sentences. S is R -refutable iff S is unsatisfiable.

Proof: Each of the rules R1 - R7 preserves validity in any model, so if S is a set of sentences and M is a model for S and $S \vdash_R E$, then $M \models E$. Consequently, if S has a model, S cannot be R -refutable.

To prove the completeness of the method, let Γ be the class of all finite sets S of universal sentences which are not R -refutable. Note that when S is any finite set of universal sentences, $S \notin \Gamma$ means that $S \vdash_R \square$. We shall show that Γ is an abstract consistency class. Now suppose that S is an unsatisfiable set of universal sentences. By the Compactness Theorem (2514) S has an unsatisfiable finite subset S_0 , and S_0 is sufficiently pure, so by Smullyan's Unifying Principle (2507) $S_0 \notin \Gamma$, so $S_0 \vdash_R \square$, so $S \vdash_R \square$.

Next we verify that Γ is an abstract consistency class.

Clearly Γ is closed under subsets.

(a') If $\square \in S$, then $S \vdash_R \square$, and if $\{A, \sim A\} \subseteq S$, then $S \vdash_R \square$ by R4, so if $S \in \Gamma$, then condition (a') is satisfied.

We check conditions (b)-(f) by the indirect method. In each case we assume that $S \in \Gamma$ and that the stated condition is false, and derive the conclusion that $S \vdash_R \square$, contradicting the assumption that $S \in \Gamma$.

- (b) We assume $\sim\sim A \in S$ and $S \cup \{A\} \vdash_R \square$. Then $S \vdash_R \square$ by $R5$.
- (c) We assume that $[A \vee B] \in S$ and $S \cup \{A\} \vdash_R \square$ and $S \cup \{B\} \vdash_R \square$. Let C_1, \dots, C_n be a refutation of $S \cup \{B\}$, and let E_1, \dots, E_m be a refutation of $S \cup \{A\}$. It is easy to establish by induction on i that for each i ($1 \leq i \leq n$), $S \vdash_R C_i$ or $S \vdash_R A \vee C_i$. Simply consider how C_i was inferred in the refutation of $S \cup \{B\}$. In case C_i is B , $A \vee C_i$ is $A \vee B$, which is in S . In case C_i was inferred by substitution, note that A is a sentence. In case C_i was inferred by Cut ($R4$), rules $R1$ and $R2$ can be used (if necessary) to infer $A \vee [M \vee N]$ from $[A \vee M] \vee [A \vee N]$. Hence $S \vdash_R \square$ or $S \vdash_R A$. In the latter case it is easy to establish that $S \vdash_R E_j$ for $1 \leq j \leq m$ by induction on j , so $S \vdash_R \square$ in this case also.
- (d) We assume that $\sim[A \vee B] \in S$ and $S \cup \{\sim A, \sim B\} \vdash_R \square$. Then $S \vdash_R \square$ by $R6$.
- (e) We assume that $\forall x A \in S$ (so $\forall x A$ is a sentence) and that there is a variable-free term t such that $S \cup \{S_t^x A\} \vdash_R \square$. Then $S \vdash_R \square$ by $R7$ and $R3$.
- (f) The condition is vacuously satisfied. Since $S \in \Gamma$, S is a set of universal sentences, so it is impossible that $\sim \forall x A \in S$. ■

3401 Theorem. Let A be any wff, let \bar{A} be its universal closure, and let M be the quantifier-free wff obtained by deleting all quantifiers from $\star \sim \bar{A}$. Let $N_1 \wedge \dots \wedge N_k$ be a conjunctive normal form of M (so each N_i is a disjunction of literals).

The following conditions are equivalent:

- (a) $\models A$.
- (b) $\{N_1, \dots, N_k\}$ has no model.
- (c) \square can be derived from $\{N_1, \dots, N_k\}$ by using rules $R1$, $R2$, $R3$, and $R4$.

Proof: Note that \bar{A} is a sentence, and $\star \sim \bar{A}$ is a rectified universal sentence, so it has $\forall x_1 \dots \forall x_n M$ as a prenex normal form, where x_1, \dots, x_n (where $n \geq 0$) are the bound variables of $\star \sim \bar{A}$. Hence

$$\vdash \star \sim \bar{A} \equiv \forall x_1 \dots \forall x_n M.$$

$$\vdash M \equiv \neg N_1 \wedge \dots \wedge N_k \quad \text{by Rule P}$$

$$\vdash \forall x_1 \dots \forall x_n M \equiv \forall x_1 \dots \forall x_n [N_1 \wedge \dots \wedge N_k] \quad \text{by the Substitutivity of Equivalence (2106)}$$

$$\vdash \forall x_1 \dots \forall x_n [N_1 \wedge \dots \wedge N_k] \equiv \neg \forall x_1 \dots \forall x_n N_1 \wedge \dots \wedge \neg \forall x_1 \dots \forall x_n N_k \quad \text{by 2129 and 2106}$$

$$\vdash \star \sim \bar{A} \equiv \neg \forall x_1 \dots \forall x_n N_1 \wedge \dots \wedge \neg \forall x_1 \dots \forall x_n N_k \quad \text{by Rule P.}$$

Hence the following conditions are equivalent:

- (a) $\models A$
- (a1) $\star \sim \bar{A}$ is unsatisfiable. by 3303
- (b1) $\forall x_1 \dots \forall x_n N_1 \wedge \dots \wedge \forall x_1 \dots \forall x_n N_k$ is unsatisfiable (by the Theorem above and 2303).
- (b2) $\{\forall x_1 \dots \forall x_n N_1, \dots, \forall x_1 \dots \forall x_n N_k\}$ has no model.
- (b) $\{N_1, \dots, N_k\}$ has no model.
- Thus (a) and (b) are equivalent.

Clearly (c) implies (b), since the rules $\mathcal{R}1 - \mathcal{R}4$ preserve validity in any model. To prove the converse, suppose that (b), and hence (b2), is true. Let S be the set $\{\forall x_1 \dots \forall x_n N_1, \dots, \forall x_1 \dots \forall x_n N_k\}$. By Theorem 3400, S is \mathcal{R} -refutable. It is not hard to see that from any \mathcal{R} -refutation of S one can construct an \mathcal{R} -refutation in which all applications of Universal Instantiation ($\mathcal{R}7$) precede all applications of the other rules. In wffs of S the scope of each negation is an atom, so Rules $\mathcal{R}5$ and $\mathcal{R}6$ can never be used. Thus the part of the refutation of S following all applications of $\mathcal{R}7$ is a refutation of $\{N_1, \dots, N_k\}$ using only $\mathcal{R}1 - \mathcal{R}4$.

EXAMPLE: Let A be $\sim \forall z[Pzx \equiv \forall u \forall v \bullet Pzu \supset \bullet Puv \supset \sim Pvz]$. Then $\sim \bar{A}$ is the wff B in Example 2 of §33, so M is

$$\begin{aligned} \sim & [[\sim Pza \vee \sim Pzu \vee \sim Puv \vee \sim Pvz] \\ & \wedge \bullet \sim [\sim Pz(gz) \vee \sim P(gz)(hz) \vee \sim P(hz)z] \vee Pza]. \end{aligned}$$

Thus A is valid iff $\{[\sim Pza \vee \sim Pzu \vee \sim Puv \vee \sim Pvz], [Pz(gz) \vee Pza], [P(gz)(hz) \vee Pza], [P(hz)z \vee Pza]\}$ has no model. We derive \square from this set as follows:

- | | | |
|------|---|------------------------|
| (1) | $\sim Pza \vee \sim Pzu \vee \sim Puv \vee \sim Pvz$ | given |
| (2) | $Pz(gz) \vee Pza$ | given |
| (3) | $P(gz)(hz) \vee Pza$ | given |
| (4) | $P(hz)z \vee Pza$ | given |
| (5) | $\sim Paa$ | Sub: 1, $\mathcal{R}2$ |
| (6) | $Pa(ga)$ | Sub: 2; Cut: 5 |
| (7) | $P(ga)(ha)$ | Sub: 3; Cut: 5 |
| (8) | $P(ha)a$ | Sub: 4; Cut: 5 |
| (9) | $\sim P(ha)a \vee \sim P(ha)a \vee \sim Pa(ga) \vee \sim P(ga)(ha)$ | Sub: 1 |
| (10) | \square | Cut: 9, 7, 6, 8 |

Thus A is valid.

If we replace Pxy by $x \in y$, then the refutable wff $\sim \forall x A$ is equivalent to $\exists x \forall z \bullet z \in x \equiv \sim \exists u \exists v \bullet z \in u \wedge u \in v \wedge v \in z$. This wff expresses an extension of Russell's Paradox due to Quine.

EXERCISES

Establish the validity of the wffs below by one of the methods discussed in this section.

X3400. $\forall x[\exists y Pxy \supset \forall y Qxy] \wedge \forall z \exists y Pzy \supset \forall y \forall x Qxy$

X3401. $\exists x \forall y [Py \equiv Px] \supset \neg \forall x Px \vee \forall x \sim Px$

X3402. $\forall y \exists w Ryw \vee \exists z \forall x [Px \supset \sim Rzx] \supset \exists x \sim Px$

X3403. $\forall x \forall y \forall z [Pxy \supset \sim Pyz] \supset \exists y \forall x \sim Pxy$

X3404. $\forall x \forall y \neg [\exists z Hxz \supset \forall z Gxz] \wedge \forall z [Gzz \supset Hzy] \supset \neg Hxy \equiv \forall z Gxz$

X3405. $\exists x \neg [Px \supset Pa] \wedge [Px \supset Pb]$

X3406. $\forall x \exists y Rxy \wedge \forall u \forall v \forall w [Ruv \wedge Rvw \supset \neg Rwu \wedge \exists s \neg Rsv \wedge Rvs] \supset \forall z Rzz$

X3407. $\forall u \forall v \forall w [Puv \vee Pvw] \supset \exists x \forall y Pxy$

Determine whether the wffs below are valid. Justify your answers by appropriate arguments.

X3408. $\forall x \forall y \neg \sim Rxy \supset \forall z \neg Rzy \supset \sim Rxz$

X3409. $\exists y \forall x \neg Px \supset \neg Py \vee Qx$

X3410. $\forall z [\exists x \forall y Pxyz \supset Qz] \supset \neg \forall z \forall y \exists x Pxyz \supset \exists z Qz$

X3411. $\forall y [\forall x Pxy \supset \forall u Quy] \supset \exists z \neg \exists v Pvz \supset \exists w \neg Pzw \vee Qwz$

§35. Herbrand's Theorem

In this section, which is adapted from [Andrews, 1981]³, we shall prove a fundamental theorem which is a consequence of the results presented by Herbrand in his thesis [Herbrand, 1930]. While this theorem involves some of the most important ideas in Herbrand's thesis, it differs from any theorem Herbrand actually stated, so at the end of this section we shall briefly indicate what Herbrand actually did.

³Adapted with permission from *Journal of the ACM*, 28 (1981), 193-214. Copyright 1981 Association for Computing Machinery, Inc.

Herbrand's Theorem is one of the most fundamental tools available for dealing with solvable cases of the decision problem⁴ and underlies many approaches to automated theorem proving. An automated theorem proving procedure called the *mating method* based on the ideas in this section is described in [Andrews, 1981]. It is closely related to the *connection method* described in [Bibel, 1987].

In this section we shall use a formulation of first-order logic in which the primitive connectives and quantifiers are \sim , \vee , \wedge , and \forall . It is not difficult to see that the results and concepts which we have obtained for \mathcal{F} also apply to this system with trivial modifications. A wff is said to be in *negation normal form (nnf)*, and is a *negation normal formula (nnf)* iff the scope of each occurrence of \sim in it is atomic. A sentence in nnf is called a *negation normal sentence (nns)*. By the methods of §14 it is easy to find for each universal wff an equivalent nnf which contains no more occurrences of atoms than the original wff. For the sake of simplicity and clarity, we shall restrict our attention to nnfs at various points in this section even though it is not essential to do so, since little real generality is lost by this restriction.

Given a wff A which we wish to show is valid, let \bar{A} be its universal closure, and let D be a negation normal form of $\star \sim \bar{A}$. Then D is a universal nns which has no model if and only if A is valid. Thus we shall concentrate on the problem of refuting universal negation normal sentences.

We first define a *compound instance (c-instance)* of a universal nnf D to be the result of replacing each wf part of D of the form $\forall x B(x)$ by $B(t_1) \wedge \dots \wedge B(t_n)$, where $n \geq 1$, and for each $i \leq n$, t_i is a closed (variable-free) term; different terms t_1, \dots, t_n (and different values of n) may be chosen for different occurrences of quantifiers in D . If $n = 1$ for each occurrence of a quantifier in D , the *c-instance* is called a *simple instance*.

To avoid any ambiguities, we define the *c*-instances of D inductively as follows:

- (a) If D is a literal, D is the only *c*-instance of D .
- (b) If D is $[D_1 * D_2]$, where $*$ is \vee or \wedge , and H_i is a *c*-instance of D_i for $i = 1, 2$, then $[H_1 * H_2]$ is a *c*-instance of D .
- (c) If D is $\forall x C(x)$, $n \geq 1$, and for each i ($1 \leq i \leq n$), t_i is a closed term and H_i is a *c*-instance of $C(t_i)$, then $[H_1 \wedge \dots \wedge H_n]$ is a *c*-instance of D .

3500 Lemma. Let D be a universal nnf, and let H be a *c*-instance of D . Then there is a conjunction M of simple instances of D such that $M \equiv H$ is tautologous; also, M and H contain exactly the same atomic wffs.

⁴See, for example, [Dreben and Goldfarb, 1979].

Proof: The proof is by induction on the number of occurrences of \vee , \wedge , and \forall in \mathbf{D} .

Case 1. \mathbf{D} is a literal. Let \mathbf{M} be \mathbf{H} .

Case 2. \mathbf{D} is $[\mathbf{D}_1 * \mathbf{D}_2]$, where $*$ is \wedge or \vee . Then \mathbf{H} has the form $[\mathbf{H}_1 * \mathbf{H}_2]$, where \mathbf{H}_i is a c -instance of \mathbf{D}_i for $i = 1, 2$. By inductive hypothesis there are conjunctions $\bigwedge_{j=1}^n \mathbf{K}_j$ and $\bigwedge_{k=1}^m \mathbf{L}_k$ of simple instances of \mathbf{D}_1 and \mathbf{D}_2 , respectively, such that $\models \mathbf{H}_1 \equiv \bigwedge_{j=1}^n \mathbf{K}_j$ and $\models \mathbf{H}_2 \equiv \bigwedge_{k=1}^m \mathbf{L}_k$. Let \mathbf{M} be $\bigwedge_{j=1}^n \bigwedge_{k=1}^m [\mathbf{K}_j * \mathbf{L}_k]$. Then \mathbf{M} is a conjunction of simple instances of \mathbf{D} , and $\models \mathbf{M} \equiv \mathbf{H}$.

Case 3. \mathbf{D} is $\forall x C(x)$. Then \mathbf{H} has the form $\mathbf{H}_1 \wedge \dots \wedge \mathbf{H}_n$, where each \mathbf{H}_i is a c -instance of $C(t_i)$ for some closed term t_i . By inductive hypothesis each \mathbf{H}_i is equivalent to a conjunction \mathbf{M}_i of simple instances of $C(t_i)$. Let \mathbf{M} be $\mathbf{M}_1 \wedge \dots \wedge \mathbf{M}_n$. It is easy to see that \mathbf{M} is a conjunction of simple instances of \mathbf{D} , and that $\models \mathbf{M} \equiv \mathbf{H}$. ■

3501 Lemma. Let \mathbf{G} and \mathbf{H} be c -instances of a universal nnf \mathbf{D} . Then there is a c -instance \mathbf{K} of \mathbf{D} such that $\models \mathbf{K} \supset [\mathbf{G} \wedge \mathbf{H}]$.

Proof: The proof is by induction on the number of occurrences of \wedge , \vee and \forall in \mathbf{D} .

Case 1. \mathbf{D} is a literal. Then $\mathbf{G} = \mathbf{D} = \mathbf{H}$, and we let $\mathbf{K} = \mathbf{D}$ also.

Case 2. \mathbf{D} has the form $\mathbf{D}_1 * \mathbf{D}_2$, where $*$ is \wedge or \vee . Then \mathbf{G} has the form $\mathbf{G}_1 * \mathbf{G}_2$, and \mathbf{H} has the form $\mathbf{H}_1 * \mathbf{H}_2$, where \mathbf{G}_i and \mathbf{H}_i are c -instances of \mathbf{D}_i such that $\models \mathbf{K}_i \supset [\mathbf{G}_i \wedge \mathbf{H}_i]$. Let $\mathbf{K} = \mathbf{K}_1 * \mathbf{K}_2$; then \mathbf{K} is a c -instance of \mathbf{D} and $\models \mathbf{K} \supset [\mathbf{G} \wedge \mathbf{H}]$.

Case 3. \mathbf{D} has the form $\forall x C(x)$. Then \mathbf{G} has the form $\mathbf{E}_1 \wedge \dots \wedge \mathbf{E}_n$ and \mathbf{H} has the form $\mathbf{E}_{n+1} \wedge \dots \wedge \mathbf{E}_m$, where for each i there is a closed term t_i such that \mathbf{E}_i is a c -instance of $C(t_i)$. Clearly we can let \mathbf{K} be $\mathbf{G} \wedge \mathbf{H}$. ■

EXAMPLE: Let D be $[\forall x Px \vee \forall y Qy]$, let G be $[Pa \vee Qb]$, H be $[Pc \vee Qd]$, and K be $[(Pa \wedge Pc) \vee (Qb \wedge Qd)]$. The wffs G , H , and K are c -instances of D , and $\models K \supset [G \wedge H]$. Note, however, that $K \equiv [G \wedge H]$ is not valid.

A *truth assignment* is an assignment of truth values (T or F) to atomic wffs. If \mathbf{G} is a quantifier-free wff and Φ is a truth assignment which assigns

truth values to all atomic wf parts of \mathbf{G} , we let $\mathcal{V}_\Phi \mathbf{G}$ denote the truth value of \mathbf{G} with respect to Φ ; this is computed in the usual way using the truth tables for the propositional connectives. We call \mathbf{G} a *truth-functional contradiction* (*t-f contradiction*) iff $\mathcal{V}_\Phi \mathbf{G} = \mathbf{F}$ for all such truth assignments Φ . We say that Φ *verifies* \mathbf{G} iff $\mathcal{V}_\Phi \mathbf{G} = \mathbf{T}$. If \mathcal{S} is a set of quantifier-free wffs and Φ is a truth assignment which verifies each member of \mathcal{S} , we say that Φ *truth-functionally satisfies* (*t-f satisfies*) \mathcal{S} .

3502 Lemma. Let \mathbf{D} be a universal nnf. If no c -instance of \mathbf{D} is a t-f contradiction, there is a truth assignment which verifies every c -instance of \mathbf{D} .

Proof: Let \mathcal{S} be the set of all c -instances of \mathbf{D} , and let $\{\mathbf{H}_1, \dots, \mathbf{H}_n\}$ be a finite subset of \mathcal{S} . Using Lemma 3501 we obtain a c -instance \mathbf{K} of \mathbf{D} such that $\models \mathbf{K} \supset [\mathbf{H}_1 \wedge \dots \wedge \mathbf{H}_n]$. Since \mathbf{K} is not a t-f contradiction, there is a truth-assignment Ψ which verifies \mathbf{K} , so Ψ t-f satisfies $\{\mathbf{H}_1, \dots, \mathbf{H}_n\}$. Thus every finite subset of \mathcal{S} is t-f satisfiable, so by the Compactness Theorem (1501) \mathcal{S} is too. ■

DEFINITIONS. The *Herbrand universe* $\mathcal{H}(\mathbf{D})$ of a sentence \mathbf{D} is the set of terms which can be constructed from the individual and function constants in \mathbf{D} (and from the individual constant c , if no other individual constant occurs in \mathbf{D}). A *compound Herbrand instance* (*cH-instance*) of a universal nns \mathbf{D} is a c -instance of \mathbf{D} in which all terms are members of $\mathcal{H}(\mathbf{D})$. A *simple Herbrand instance* of \mathbf{D} is a *cH-instance* which is a simple instance of \mathbf{D} .

3503 Herbrand's Theorem. Let \mathbf{A} be a wff, let $\overline{\mathbf{A}}$ be its universal closure, and let \mathbf{D} be a negation normal form of the Skolemization $\star \sim \overline{\mathbf{A}}$ of $\sim \overline{\mathbf{A}}$. Then \mathbf{A} is valid if and only if \mathbf{D} has a compound Herbrand instance which is a truth-functional contradiction.

Proof: We may assume that we are using a formulation of first-order logic in which there are no individual or function constants except those from which the terms in $\mathcal{H}(\mathbf{D})$ are constructed. Thus all c -instances of \mathbf{D} are *cH*-instances.

Clearly $\models \mathbf{A}$ iff \mathbf{D} has no model. Thus we shall show that \mathbf{D} has no model iff \mathbf{D} has a contradictory c -instance.

It is easy to see by induction on the number of occurrences of \wedge , \vee , and \forall in \mathbf{D} , that if \mathbf{D} is any universal nns, and \mathbf{G} is any c -instance of \mathbf{D} , then $\models \mathbf{D} \supset \mathbf{G}$. Hence if \mathbf{G} is t-f contradictory, it is false in any model, so \mathbf{D} is

too, so \mathbf{D} has no model.

For the proof in the other direction, suppose \mathbf{D} has no $t\text{-}f$ contradictory c -instance. Then by Lemma 3502 there is an assignment Φ of truth values to the atoms which occur in c -instances of \mathbf{D} which verifies every c -instance of \mathbf{D} .

We next construct a model for \mathbf{D} . We let the domain of individuals of our model be the Herbrand universe of \mathbf{D} . As in the proof of 2507, we interpret the individual constants and function symbols so that if P is any n -ary predicate and t_1, \dots, t_n are individuals, then Pt_1, \dots, t_n is true in the interpretation iff $\Phi(Pt_1, \dots, t_n) = T$. This specifies an interpretation \mathcal{M} for our system of first-order logic.

Next, we show, by induction on the number of occurrences of \wedge , \vee , and \forall in \mathbf{B} , that if \mathbf{B} is any universal nns such that Φ verifies every c -instance of \mathbf{B} , then $\mathcal{M} \models \mathbf{B}$.

Case 1. \mathbf{B} is a literal. Then \mathbf{B} is the only c -instance of \mathbf{B} , and $\mathcal{V}^{\mathcal{M}}\mathbf{B} = \mathcal{V}_{\Phi}\mathbf{B}$ by the interpretation of predicates.

Case 2. \mathbf{B} has the form $\mathbf{B}_1 \wedge \mathbf{B}_2$. For $i = 1, 2$, let \mathbf{G}_i be any c -instance of \mathbf{B}_i . Then $\mathbf{G}_1 \wedge \mathbf{G}_2$ is a c -instance of \mathbf{B} , so $\mathcal{V}_{\Phi}[\mathbf{G}_1 \wedge \mathbf{G}_2] = T$, so $\mathcal{V}_{\Phi}\mathbf{G}_i = T$. Therefore, $\mathcal{M} \models \mathbf{B}_i$ by inductive hypothesis, so $\mathcal{M} \models \mathbf{B}$.

Case 3. \mathbf{B} has the form $[\mathbf{B}_1 \vee \mathbf{B}_2]$. Suppose there are c -instances \mathbf{G}_1 of \mathbf{B}_1 and \mathbf{G}_2 of \mathbf{B}_2 such that $\mathcal{V}_{\Phi}\mathbf{G}_1 = F$ and $\mathcal{V}_{\Phi}\mathbf{G}_2 = F$. Then $\mathcal{V}_{\Phi}[\mathbf{G}_1 \vee \mathbf{G}_2] = F$, which contradicts our assumption about \mathbf{B} , since $[\mathbf{G}_1 \vee \mathbf{G}_2]$ is a c -instance of \mathbf{B} . Therefore for $i = 1$ or $i = 2$, Φ verifies every c -instance of \mathbf{B}_i , so $\mathcal{M} \models \mathbf{B}_i$ by inductive hypothesis, so $\mathcal{M} \models \mathbf{B}$.

Case 4. \mathbf{B} has the form $\forall x C(x)$. Let t be any closed term. Every c -instance of $C(t)$ is a c -instance of \mathbf{B} , and so is verified by Φ , so $\mathcal{M} \models C(t)$ by inductive hypothesis. Therefore $\mathcal{M} \models \forall x C(x)$, since every individual of \mathcal{M} is the denotation of some closed term.

(In case the argument in the preceding sentence seems too informal, we can make it more rigorous as follows. Let φ be any assignment, and let ψ be any assignment which agrees with φ off x . There is some closed term t such that $\psi x = t = \mathcal{V}^{\mathcal{M}}t$, so by the Corollary to the Substitution-Value Theorem (2302), $\mathcal{V}_{\psi}^{\mathcal{M}}C(x) = \mathcal{V}_{\varphi}^{\mathcal{M}}S_t^x C(x) = \mathcal{V}_{\varphi}^{\mathcal{M}}C(t) = T$. Thus $\mathcal{V}_{\varphi}^{\mathcal{M}}\forall x C(x) = T$, and φ is arbitrary, so $\mathcal{M} \models \forall x C(x)$.)

This completes the inductive argument, and establishes that $\mathcal{M} \models \mathbf{D}$, so \mathbf{D} has a model. ■

Alternative formulations of Herbrand's Theorem are provided by the following corollaries:

3504 Corollary. Let \mathbf{A} be a wff, let $\overline{\mathbf{A}}$ be its universal closure, and let \mathbf{D} be a negation normal form of the Skolemization $\star \sim \overline{\mathbf{A}}$ of $\sim \overline{\mathbf{A}}$. Then \mathbf{A} is valid if and only if some conjunction of simple Herbrand instances of \mathbf{D} is a truth-functional contradiction.

Proof: By 3500 every cH -instance of \mathbf{D} is truth-functionally equivalent to a conjunction of simple Herbrand instances, and by 3501 every conjunction of simple Herbrand instances of \mathbf{D} is truth-functionally implied by some cH -instance, so the corollary follows from 3503. ■

3505 Corollary. Let \mathbf{B} be a sentence, and let \mathbf{D} be a negation normal form of the Skolemization $\star \mathbf{B}$ of \mathbf{B} . Then \mathbf{B} is satisfiable if and only if the set of simple Herbrand instances of \mathbf{D} is t-f satisfiable.

Proof: Let \mathbf{A} be $\sim \mathbf{B}$. Then \mathbf{D} is a nnf of $\star \sim \overline{\mathbf{A}}$, since $\sim \overline{\mathbf{A}}$ is $\sim \sim \mathbf{B}$. Since \mathbf{B} is satisfiable iff \mathbf{A} is not valid, by 3504 \mathbf{B} is satisfiable iff every conjunction of simple Herbrand instances of \mathbf{D} is t-f satisfiable. The corollary follows from this by the Compactness Theorem (1501). ■

EXAMPLE: Suppose we wish to establish the validity of the following wff:

$$(A) \exists x \forall y [Px \equiv Py] \supset [\exists x Px \equiv \forall y Py]$$

We negate it, and eliminate \equiv and \supset to obtain:

$$(B) \exists x \forall y [(\sim Px \vee Py) \wedge (\sim Py \vee Px)] \\ \wedge ([\exists x Px \wedge \exists y \sim Py] \vee [\forall y Py \wedge \forall x \sim Px])$$

Next we Skolemize it by the methods of §33 to obtain the following nns:

$$(D) \forall y [(\sim Pd \vee Pe) \wedge (\sim Pe \vee Pd)] \\ \wedge ([\forall z Pz \wedge \forall x \sim Px])$$

It is convenient to display this in the 2-dimensional format discussed in §14:

$$(D') \left[\begin{array}{c} \forall y \left[\begin{array}{c} \sim P_c \vee P_y \\ \sim P_y \vee P_c \end{array} \right] \\ \left[\begin{array}{c} P_d \\ \sim P_e \end{array} \right] \vee \left[\begin{array}{c} \forall z P_z \\ \forall x \sim P_x \end{array} \right] \end{array} \right]$$

The following is a *c*-instance of (D):

$$(E) \left[\begin{array}{ccc} \sim P_c & \vee & P_d \\ \sim P_d & \vee & P_c \\ \sim P_c & \vee & P_e \\ \sim P_e & \vee & P_c \\ \left[\begin{array}{c} P_d \\ \sim P_e \end{array} \right] \vee \left[\begin{array}{c} P_c \\ \sim P_d \end{array} \right] \end{array} \right]$$

It can be easily be seen by the methods of §14 (especially 1408) that (E) is contradictory, so by 3503 (A) is valid.

Since a variety of results in logic are sometimes called Herbrand's theorem, we conclude this section with a brief and rather imprecise description of the main results which Herbrand actually presented in this thesis.

Herbrand introduced a cut-free system of first-order logic which we shall call \mathcal{F}_H . \mathcal{F}_H had all quantifier-free tautologies as axioms, and had as rules of inference universal and existential generalization, anti-prenex rules for pushing in quantifiers, the rule of alphabetic change of bound variables, and a rule of simplification (replace $D \vee D$ by D). Modus ponens was not a rule of inference of \mathcal{F}_H .

He also considered a system of first-order logic which did have modus ponens as a primitive rule of inference, and which we may pretend was the system \mathcal{F} of Chapter 2. (Systems of this sort are called *Hilbert-style* systems.)

He also showed how to associate with a wff A of first-order logic certain quantifier-free wffs which we shall call *Herbrand expansions* of A , which were equivalent to negations of the *cH*-instances of the wff D mentioned in 3503. Let us say that a wff A has the *Herbrand property* iff some Herbrand expansion of A is tautologous.

Herbrand's Theorem consisted of the following claims about any wff A of first-order logic:

- (1) If $\vdash_{\mathcal{F}_H} A$, then $\vdash_{\mathcal{F}} A$.
- (2) If $\vdash_{\mathcal{F}} A$, then A has the Herbrand property.
- (3) If A has the Herbrand property, then $\vdash_{\mathcal{F}_H} A$.

Thus Herbrand claimed that the following are equivalent:

- (a) $\vdash_{\mathcal{F}_H} \mathbf{A}$
- (b) $\vdash_{\mathcal{F}} \mathbf{A}$
- (c) \mathbf{A} has the Herbrand property.

If we add the soundness (2303) and completeness (2510) of \mathcal{F} to the claims above, it is natural to add to this list of equivalent statements the following:

- (d) \mathbf{A} is valid.

Various versions of the theorem that (c) and (d) are equivalent, such as our theorem 3503, are commonly called Herbrand's Theorem in the literature of logic. Notice, however, that Herbrand's version of his theorem was a purely syntactic result. His proofs were purely syntactic too, and it is much more difficult to prove these results by purely syntactic arguments than it is to prove them with the aid of semantic as well as syntactic concepts. Actually, Herbrand's proof of (2) contained a basic error⁵ [Dreben *et al.*, 1963], which was finally corrected by Dreben and Denton [Dreben and Denton, 1966].

Note that the full Herbrand theorem implies that modus ponens is a derived rule of inference of the system \mathcal{F}_H , a result that is closely related to Gentzen's Cut-Elimination Theorem (3102).

§36. Unification

Most systematic theorem-proving procedures involve making substitutions for variables in such a way that certain expressions become the same. In this case we say that they are *unified*. For example, when using the methods of §34, the purpose of applying the substitution rule is to prepare for applications of Simplification or Cut by unifying certain expressions. Similarly, when using the methods of §35, one must find a contradictory Herbrand

⁵ It is reported in [van Heijenoort, 1967, p. 525] that Kurt Gödel discovered an essential gap in Herbrand's proof in the early 1940's, but he never published anything on the subject. In 1962 Andrews discovered a gap in Herbrand's proof while examining it to see if it could be extended to higher-order logic. He communicated this to Dreben, who was an authority on Herbrand's work. After some discussion Dreben became convinced that there was indeed a gap in the proof, and set about repairing it. Instead, he found a counterexample to one of Herbrand's lemmas. An attempt was made to replace that lemma by a weaker lemma which would still suffice to prove Herbrand's Theorem, but Aanderaa found a counterexample to this.

instance of a sentence, and (as one sees from 1408) this requires substituting terms for the quantified variables in such a way that certain expressions become unified.

When trying to prove a theorem by such methods, one often proceeds in steps, and each step involves unifying additional expressions. Since a set of expressions may have many unifiers, it is advisable to use at each stage a unifier which is as general as possible, i.e., one which leaves as much freedom as possible for imposing additional constraints on the terms to be finally substituted for the variables.

Before discussing an example, let us clarify some terminology. In this section we shall use the word *expression* to mean an atomic formula or a term from a formulation of first order logic in which the only variables are individual variables. Also, by a *substitution* we shall mean a substitution for variables. (Review the definitions at the end of §10.) If \mathcal{W} is a set of expressions and θ is a substitution, we define $\theta\mathcal{W}$ to be $\{\theta\mathbf{X} \mid \mathbf{X} \in \mathcal{W}\}$. If $\theta\mathcal{W}$ is a unit set (set with exactly one member), we say that θ *unifies* \mathcal{W} , and is a *unifier of* \mathcal{W} . This means, of course, that $\theta\mathbf{X} = \theta\mathbf{Y}$ for all \mathbf{X} and \mathbf{Y} in \mathcal{W} .

EXAMPLE: Let $A = \bar{Q}^2 xy$, $B = \bar{Q}^2 z \bar{h}^1 z$, $C = \bar{Q}^2 x \bar{h}^1 \bar{a}$, and $D = \bar{Q}^2 \bar{g}^2 uvv$, where \bar{Q}^2 , \bar{g}^2 , and \bar{h}^1 are predicate and function constants, and \bar{a} is an individual constant. (We shall henceforth omit the superscripts.) Suppose we wish to simultaneously unify $\{A, B\}$ and $\{C, D\}$, i.e., find a substitution θ such that $\theta A = \theta B$ and $\theta C = \theta D$. We might do this by finding a substitution τ which unifies $\{A, B\}$ and then finding a substitution σ which unifies $\{\tau C, \tau D\}$, and letting $\theta = \sigma \circ \tau$. In this case $\theta A = (\sigma \circ \tau)A = \sigma(\tau A) = \sigma(\tau B) = \theta B$ and $\theta C = \sigma(\tau C) = \sigma(\tau D) = \theta D$. Let $\tau_1 = S_{\bar{a} \bar{h} \bar{a} \bar{a}}^{x y z}$. We might happen to notice that τ_1 unifies $\{A, B\}$, and proceed to try to unify $\{\tau_1 C, \tau_1 D\}$, which is $\{\bar{Q} \bar{a} \bar{h} \bar{a}, \bar{Q} \bar{g} u v v\}$, but this set is not unifiable. Let $\tau = S_{z \bar{h} z}^{x y}$ and $\sigma = S_{\bar{h} \bar{a} \bar{g} u \bar{h} \bar{a}}^v z$. τ also unifies $\{A, B\}$, and $\{\tau C, \tau D\}$, which is $\{\bar{Q} z \bar{h} \bar{a}, \bar{Q} \bar{g} u v v\}$ has a unifier, namely σ . It can be seen that $\sigma \circ \tau = S_{\bar{h} \bar{a} \bar{g} u \bar{h} \bar{a} \bar{h} \bar{g} u \bar{h} \bar{a} \bar{g} u \bar{h} \bar{a}}^{v x y z}$ (for example, $(\sigma \circ \tau)y = \sigma(\tau y) = \sigma(\bar{h} z) = \bar{h} \sigma z = \bar{h} g u \bar{h} \bar{a}$), and that $\sigma \circ \tau$ unifies both $\{A, B\}$ and $\{C, D\}$. Indeed, $(\sigma \circ \tau)A = \bar{Q} \bar{g} u \bar{h} \bar{a} \bar{h} \bar{g} u \bar{h} \bar{a} = (\sigma \circ \tau)B$ and $(\sigma \circ \tau)C = \bar{Q} \bar{g} u \bar{h} \bar{a} \bar{h} \bar{a} = (\sigma \circ \tau)D$.

Clearly τ is a more general substitution than τ_1 . Next we make this notion precise.

DEFINITION. Let τ_0 and τ_1 be substitutions. We write $\tau_0 \leq \tau_1$, and say that τ_0 is *less specified than* τ_1 , or that τ_0 is *more general than* τ_1 , iff there is a

substitution ρ such that $\rho \circ \tau_0 = \tau_1$.

In the example above, $S_a^z \circ \tau = \tau_1$, so $\tau \leq \tau_1$.

It is easy to see that the relation \leq is reflexive (X3600) and transitive (X3601).

DEFINITION. Let \mathcal{W} be a set of expressions and θ be substitution. θ is a *most general unifier (mgu)* of \mathcal{W} iff θ unifies \mathcal{W} and for any unifier τ of \mathcal{W} , $\theta \leq \tau$.

DEFINITION. Let τ_1 and τ_2 be substitutions. We write $\tau_1 \approx \tau_2$ iff $\tau_1 \leq \tau_2$ and $\tau_2 \leq \tau_1$.

EXAMPLE: Let $\tau_1 = S_{hx}^w y$, $\tau_2 = S_{hy}^w x$, and $\rho = S_y^x y$. Then $\rho \circ \tau_1 = \tau_2$ and $\rho \circ \tau_2 = \tau_1$, so $\tau_1 \approx \tau_2$.

EXAMPLE: Let \bar{k} be a ternary function constant, let $\tau_1 = S_{\bar{k}xyz}^w x y z$, $\tau_2 = S_{\bar{k}yzx}^w x y z$, $\sigma_1 = S_y^x z y$, and $\sigma_2 = S_z^x y z$. Then $\sigma_1 \circ \tau_1 = \tau_2$ and $\sigma_2 \circ \tau_2 = \tau_1$, so $\tau_1 \approx \tau_2$.

It is easy to see that \approx is an equivalence relation (X3602). It can also be seen that substitutions which are equivalent in this sense differ only by alphabetic changes of variables. The next proposition shows that most general unifiers of sets of expressions of first-order logic are essentially unique.

3600 Proposition. If \mathcal{W} is a set of expressions and θ and τ are most general unifiers of \mathcal{W} , then $\theta \approx \tau$.

Proof: Under the hypotheses θ and τ both unify \mathcal{W} , and $\theta \leq \tau$ and $\tau \leq \theta$, so $\theta \approx \tau$. ■

3601 Unification Theorem [Robinson, 1965]. There is an algorithm which can be applied to an arbitrary nonempty finite set \mathcal{W} of expressions to determine whether \mathcal{W} is unifiable. The algorithm always terminates, and produces a most general unifier of \mathcal{W} if \mathcal{W} is unifiable.

Proof: If \mathcal{W} is a finite set of expressions with at least two members, define the *disagreement set* $D(\mathcal{W})$ of \mathcal{W} as follows. The expressions in \mathcal{W} are finite sequences of symbols. Let n be the first position at which not all expressions in \mathcal{W} have the same symbol, and let $D(\mathcal{W})$ be the set of expressions \mathbf{B} such that there is an expression \mathbf{A} in \mathcal{W} such that \mathbf{B} is a sub-expression of \mathbf{A} (expression which is a consecutive subsequence of \mathbf{A}) and the first symbol of \mathbf{B} is the n th symbol of \mathbf{A} .

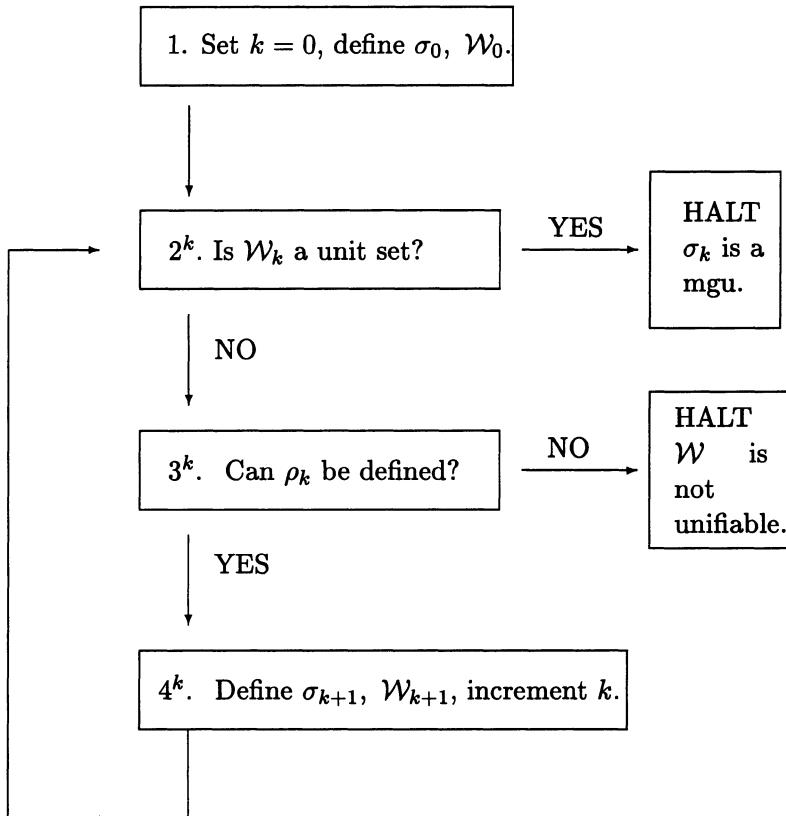


Figure 3.5: Flow chart for the unification algorithm

EXAMPLE: Let $\mathcal{W} = \{\bar{Q}^2x\bar{h}^1\bar{a}, \bar{Q}^2\bar{g}^2uvv, \bar{Q}^2\bar{g}^2\bar{h}^1x\bar{g}^2u\bar{h}^1yx\}$. Then $n = 2$ and $D(\mathcal{W}) = \{x, \bar{g}^2uv, \bar{g}^2\bar{h}^1x\bar{g}^2u\bar{h}^1y\}$.

We next describe how the algorithm, known as the *Unification Algorithm*, works when applied to a nonempty finite set \mathcal{W} of expressions. (See Figure 3.5.)

Step 1. Set $k = 0$, $\mathcal{W}_0 = \mathcal{W}$, and let σ_0 be the identity substitution. Go to Step 2⁰.

Step 2^k. If \mathcal{W}_k is a unit set, halt; σ_k is a most general unifier of \mathcal{W} . Otherwise, go to Step 3^k.

Step 3^k. If the disagreement set $D(\mathcal{W}_k)$ contains a variable v_k and a

term t_k such that v_k does not occur in t_k , choose such a v_k and t_k in some systematic way, let $\rho_k = S_{t_k}^{v_k}$, and go to Step 4^k . Otherwise halt; \mathcal{W} is not unifiable.

Step 4^k . Let $\sigma_{k+1} = \rho_k \circ \sigma_k$ and $\mathcal{W}_{k+1} = \rho_k \mathcal{W}_k$. Set k to $k + 1$, and then go to Step 2^k .

It is easy to see by induction on k that for each k , $\mathcal{W}_k = \sigma_k \mathcal{W}$, that $\sigma_{k+1} = \rho_k \circ \rho_{k-1} \circ \dots \circ \rho_2 \circ \rho_1 \circ \rho_0$, and that none of the variables v_0, \dots, v_k occurs in any expression in \mathcal{W}_{k+1} . Since v_k is chosen from some expression \mathcal{W}_k in Step 3^k , the variables v_0, v_1, v_2, \dots are clearly all distinct from one another and all occur in expressions of \mathcal{W} . There are only finitely many such variables, so the algorithm clearly terminates.

If it terminates at Step 2^k , then σ_k unifies \mathcal{W} . Therefore, if \mathcal{W} is not unifiable, the algorithm cannot terminate at Step 2^k for any k , so it must terminate at some Step 3^k .

Next consider the case where \mathcal{W} is unifiable. Let θ be any unifier of \mathcal{W} . We shall first show that whenever σ_k is defined and $\sigma_k \leq \theta$, the algorithm does not terminate at Step 3^k . If \mathcal{W}_k is a unit set, the algorithm terminates at Step 2^k . Otherwise, it reaches Step 3^k , and $D(\mathcal{W}_k)$ has at least two members. Since $\sigma_k \leq \theta$, there is a substitution τ such that $\tau \circ \sigma_k = \theta$. $\tau \mathcal{W}_k = \tau(\sigma_k \mathcal{W}) = (\tau \circ \sigma_k) \mathcal{W} = \theta \mathcal{W}$, which is a unit set, so τ unifies \mathcal{W}_k , and hence $D(\mathcal{W}_k)$. If the first symbol of each expression in $D(\mathcal{W}_k)$ were a constant, $D(\mathcal{W}_k)$ would not be unifiable, since (by the definition of a disagreement set) these constants could not all be the same. Therefore the first symbol of some expression X in $D(\mathcal{W}_k)$ is a variable, and therefore X is a variable, which we shall call v_k . Let t_k be any other expression in $D(\mathcal{W}_k)$. Since $\tau v_k = \tau t_k$ but $v_k \neq t_k$, it is clear that v_k cannot occur in t_k ; otherwise τv_k would be a proper subsequence of τt_k , which is τv_k . Therefore the algorithm does not terminate at Step 3^k .

Next we show by induction on k that $\sigma_k \leq \theta$ for each k such that σ_k is defined. This is clear for $k = 0$, since $\theta \circ \sigma_0 = \theta$. For the induction step suppose $\sigma_k \leq \theta$ and σ_k is defined. If the algorithm terminates at Step 2^k , then σ_{k+1} is not defined. Otherwise, it proceeds to Step 3^k , where it does not terminate (as shown above), so $v_k, t_k, \rho_k, \sigma_{k+1}$, and \mathcal{W}_{k+1} are defined as in Steps 3^k and 4^k . We are given that there is a substitution τ such that $\tau \circ \sigma_k = \theta$, and (as shown above) $\tau v_k = \tau t_k$.

Define the substitution τ' so that $\tau' v_k = v_k$, and $\tau' v = \tau v$ for every variable v distinct from v_k . We show that $\tau' \circ \rho_k = \tau$ by showing that these substitutions agree on all variables. $(\tau' \circ \rho_k) v_k = \tau'(\rho_k v_k) = \tau' t_k = \tau t_k$ (since v_k does not occur in t_k , and τ' agrees with τ on all other variables) $= \tau v_k$. If v is a variable distinct from v_k , then $(\tau' \circ \rho_k) v = \tau'(\rho_k v) = \tau' v =$

τv . Thus $\tau' \circ \rho_k = \tau$.

Now $\tau' \circ \sigma_{k+1} = \tau' \circ (\rho_k \circ \sigma_k)$ (by the definition of σ_{k+1}) $= (\tau' \circ \rho_k) \circ \sigma_k = \tau \circ \sigma_k = \theta$, so $\sigma_{k+1} \leq \theta$, as required. This completes the inductive argument.

Since $\sigma_k \leq \theta$ whenever σ_k is defined, and therefore the algorithm cannot terminate at any Step 3^k , it must terminate at some Step 2^K . Since we have shown that $\sigma_K \leq \theta$ for an arbitrary unifier θ of \mathcal{W} , σ_K is a most general unifier of \mathcal{W} . ■

There is much more to be said about unification. One can perform unification modulo certain theories (such as the associativity or commutativity of certain functions) or in a richer language (such as type theory). One can also seek more efficient unification algorithms. We refer the reader to [Baader and Siekmann, 1994], [Baader and Snyder, 2000], [Degtyarev and Voronkov, 1998], [Gallier and Snyder, 1989], [Gallier *et al.*, 1992], [Huet and Oppen, 1980], [Jouannaud and Kirchner, 1991], [Paterson and Wegman, 1978], [Siekmann, 1984], [Siekmann, 1989], [Snyder, 1991], and [van Vaaalen, 1975] for more details and references.

EXERCISES

X3600. Show that the relation \leq is reflexive.

X3601. Show that the relation \leq is transitive.

X3602. Show that \approx is an equivalence relation on the set of substitutions.

Chapter 4

Further Topics in First-Order Logic

§40. Duality

We introduce this subject with a parable about two scholars, named Oren and Nero, who were visiting an archaeologist and were shown a recently discovered tablet, the contents of which are reproduced in Figure 4.1. They soon realized that the figures on the tablet were truth tables, and they set about translating them into more familiar notations. (Before proceeding further, the reader is advised to do this for himself for at least a few of the tables.) Oren produced the translation in Figure 4.2, and Nero produced that in Figure 4.3. As they started to show their translations to the archaeologist, Nero modestly remarked “That was really quite easy, as soon as I realized that \oplus denoted conjunction”. “But you’re wrong!” exclaimed Oren. “ \oplus denoted disjunction!” They argued for some time, and neither was able to persuade the other that he was wrong. All they could agree on was that \ominus denoted negation.

Then the archaeologist showed them an inscription which had been found on another tablet. “We’ve figured out how to translate most of the language they apparently used in everyday affairs”, he said, “but this inscription seems to be in a mixture of two languages. Here’s how the inscription looks under the translation we’ve achieved so far:”

‘The value of $(\boxplus x)A(x)$ is $\nabla\nabla$ iff there is an n such that the value of $A(n)$ is $\nabla\nabla$.’

Oren and Nero recognized that this was another fragment of a text on logic,

\circlearrowleft	Δ	∇	\square	∇	Δ
Δ	Δ	Δ	∇	∇	∇
∇	Δ	Δ	Δ	∇	∇
\oplus	Δ	∇	\boxplus	∇	Δ
Δ	Δ	Δ	∇	∇	∇
∇	Δ	∇	Δ	∇	Δ
\otimes	Δ	∇	\boxtimes	∇	Δ
Δ	Δ	Δ	∇	∇	∇
∇	∇	Δ	Δ	∇	∇
\oslash	Δ	∇	\boxslash	∇	Δ
Δ	Δ	∇	∇	∇	Δ
∇	Δ	Δ	Δ	∇	∇
\otimes	Δ	∇	\boxtimes	∇	Δ
Δ	Δ	∇	∇	∇	Δ
∇	∇	Δ	Δ	Δ	∇
\odot	Δ	∇	\boxdot	∇	Δ
Δ	∇	∇	∇	Δ	∇
∇	∇	Δ	Δ	Δ	∇
Θ				Θ	
Δ			Δ		
∇			∇		

Figure 4.1: The tablet

and each soon produced a translation of it. Oren's translation was:

'The value of $\forall x A(x)$ is falsehood iff there is an n such that the value of $A(n)$ is falsehood.'

Nero's translation was:

'The value of $\exists x A(x)$ is truth iff there is an n such that the value of $A(n)$ is truth.'

Again, they could find no way to resolve their disagreement.

t	T	F
T	T	T
F	T	T

f	F	T
F	F	F
T	F	F

f	F	T
F	F	F
T	F	F

t	T	F
T	T	T
F	T	T

\vee	T	F
T	T	T
F	T	F

\wedge	F	T
F	F	F
T	F	T

\wedge	F	T
F	F	F
T	F	T

\vee	T	F
T	T	T
F	T	F

c	T	F
T	T	T
F	F	T

$\not\exists$	F	T
F	F	F
T	T	F

$\not\exists$	F	T
F	F	F
T	T	F

c	T	F
T	T	T
F	F	T

\supset	T	F
T	T	F
F	T	T

$\not\subset$	F	T
F	F	T
T	F	F

$\not\subset$	F	T
F	F	T
T	T	F

\supset	T	F
T	T	F
F	T	T

\equiv	T	F
T	T	F
F	F	T

$\not\equiv$	F	T
F	F	T
T	T	F

$\not\equiv$	F	T
F	F	T
T	T	F

\equiv	T	F
T	T	F
F	F	T

\downarrow	T	F
T	F	F
F	F	T

\mid	F	T
F	T	T
T	T	F

\mid	F	T
F	T	T
T	T	F

\downarrow	T	F
T	F	F
F	F	T

	\sim
T	F
F	T

	\sim
F	T
T	F

Figure 4.2: Oren's translation of the tablet

Figure 4.3: Nero's translation of the tablet

"There's one more fragment of a tablet you might be interested in", said the archeologist. "We managed to translate the first word on it as 'AXIOMS', and there's a line under that which looks like this."

$$(\oplus p)[p \oplus \ominus p].$$

"That settles it!" exclaimed Oren. "That translates to ' $\forall p[p \vee \sim p]$ ', which is a fine axiom. Clearly my method of translation is correct." "Not at all", replied Nero. "It translates to ' $\exists p[p \wedge \sim p]$ '. The people who wrote these

tablets were known throughout the ancient world as deceitful, treacherous, inveterate liars. Apparently they even axiomatized their lies!"

We leave it to the reader to decide what the moral of this story is.

Clearly there is a symmetry, or *duality*, in logic induced by systematically interchanging truth and falsehood. We shall show how this duality can be used. The ideas in this section apply to any sound and complete formulation of first-order logic. We shall present them in terms of \mathcal{F} , but they apply equally well to a system in which both \forall and \exists , and all binary propositional connectives, are primitive symbols.

When \subset , $\not\subset$, $\not\subseteq$, $\not\equiv$, \downarrow , and $|$ occur in abbreviations of wffs of \mathcal{F} , they are to be regarded as having been introduced by the following definitions, which we hereby add to the definitions of \wedge , \supset , and \equiv in §10, and of \exists in §20:

$[A \subset B]$	stands for	$[A \vee \sim B]$.
$[A \not\subset B]$	stands for	$\sim [\sim A \vee B]$.
$[A \not\subseteq B]$	stands for	$\sim [A \vee \sim B]$.
$[A \not\equiv B]$	stands for	$\sim [A \equiv B]$.
$[A \downarrow B]$	stands for	$\sim [A \vee B]$.
$[A B]$	stands for	$\sim [A \wedge B]$.

Definition. Let A be a wff of \mathcal{F} . A wff B is a *dual* of A iff there is an abbreviation C for A such that B is the result of interchanging \forall with \exists , \vee with \wedge , \supset with $\not\subseteq$, \subset with $\not\subset$, \equiv with $\not\equiv$, and \downarrow with $|$ everywhere in C (See Figure 4.4.) The *principal dual* A^d of A is the result of making these interchanges in the unabbreviated form of A .

\forall	\exists
\wedge	\vee
\supset	$\not\subseteq$
\subset	$\not\subset$
\equiv	$\not\equiv$
\downarrow	$ $

Figure 4.4: Dual Pairs

4000 Proposition. If B and D are duals of A , then $\vdash B \equiv D$.

Proof: We show that if B is any dual of A , then $\vdash B \equiv A^d$. Clearly this will suffice to prove the proposition. We show, by induction on the construction

of \mathbf{D} , that if \mathbf{D} is any abbreviation for a wff \mathbf{D}_0 of \mathcal{F} , and \mathbf{D}' is the result of interchanging \forall with \exists , \vee with \wedge , \supset with \subset , etc., in \mathbf{D} , then $\vdash \mathbf{D}' \equiv \mathbf{D}_0^d$.

Of course, this is trivially true if \mathbf{D} is atomic. Now consider the following cases:

- (a) \mathbf{D} is $\forall \mathbf{x} \mathbf{M}$. Then $\mathbf{D}' = \exists \mathbf{x} \mathbf{M}'$, but $\vdash \mathbf{M}' \equiv \mathbf{M}_0^d$ by inductive hypothesis, so $\vdash \mathbf{D}' \equiv \exists \mathbf{x} \mathbf{M}_0^d$. Also $\mathbf{D}_0^d = (\forall \mathbf{x} \mathbf{M}_0)^d = \exists \mathbf{x} \mathbf{M}_0^d$, so $\vdash \mathbf{D}' \equiv \mathbf{D}_0^d$.
- (b) \mathbf{D} is $\exists \mathbf{x} \mathbf{M}$. Then $\mathbf{D}' = \forall \mathbf{x} \mathbf{M}'$, but $\vdash \mathbf{M}' \equiv \mathbf{M}_0^d$ by inductive hypothesis, so $\vdash \mathbf{D}' \equiv \forall \mathbf{x} \mathbf{M}_0^d$. Also $\mathbf{D}_0^d = (\sim \forall \mathbf{x} \sim \mathbf{M}_0)^d = \sim \exists \mathbf{x} \sim \mathbf{M}_0^d$, so $\vdash \mathbf{D}' \equiv \mathbf{D}_0^d$.
- (c) \mathbf{D} is $[\mathbf{M} \vee \mathbf{N}]$. Then $\mathbf{D}' = [\mathbf{M}' \wedge \mathbf{N}']$, but $\vdash \mathbf{M}' \equiv \mathbf{M}_0^d$ and $\vdash \mathbf{N}' \equiv \mathbf{N}_0^d$ by inductive hypothesis, so $\vdash \mathbf{D}' \equiv [\mathbf{M}_0^d \wedge \mathbf{N}_0^d]$. Also $\mathbf{D}_0^d = [\mathbf{M}_0 \vee \mathbf{N}_0]^d = [\mathbf{M}_0^d \wedge \mathbf{N}_0^d]$, so $\vdash \mathbf{D}' \equiv \mathbf{D}_0^d$.
- (d) \mathbf{D} is $[\mathbf{M} \wedge \mathbf{N}]$. Then $\mathbf{D}' = [\mathbf{M}' \vee \mathbf{N}']$, but $\vdash \mathbf{M}' \equiv \mathbf{M}_0^d$ and $\vdash \mathbf{N}' \equiv \mathbf{N}_0^d$ by inductive hypothesis, so $\vdash \mathbf{D}' \equiv [\mathbf{M}_0^d \vee \mathbf{N}_0^d]$. Also $\mathbf{D}_0^d = (\sim [\sim \mathbf{M}_0 \vee \sim \mathbf{N}_0])^d = \sim [\sim \mathbf{M}_0^d \wedge \sim \mathbf{N}_0^d] = \sim \sim [\sim \sim \mathbf{M}_0^d \vee \sim \sim \mathbf{N}_0^d]$, so $\vdash \mathbf{D}' \equiv \mathbf{D}_0^d$.

Since the pattern of proof is now clear, we leave it to the reader to check the remaining cases (Exercise X4000). Note that there is some purpose to this exercise, since it should be verified that the duals of the remaining connectives are defined correctly. ■

4001 Corollary. For any wff \mathbf{A} of \mathcal{F} , $\vdash \mathbf{A}^{dd} \equiv \mathbf{A}$.

Proof: It is easy to see that \mathbf{A} is a dual of \mathbf{A}^d , so $\vdash \mathbf{A}^{dd} \equiv \mathbf{A}$ by 4000. ■

REMARK. Before looking at the next theorem, the reader should test his understanding of the basic idea of duality by trying to figure out what the theorem should say. The theorem has the form $\mathcal{V}_\varphi^\mathcal{M} \mathbf{A}^d = ?$, and essentially says that one can compute $\mathcal{V}_\varphi^\mathcal{M} \mathbf{A}^d$ if one knows $\mathcal{V}_\varphi^\mathcal{N} \mathbf{A}$ for an appropriate interpretation \mathcal{N} .

4002 Theorem. Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$ be any interpretation. Define \mathcal{J}' to agree with \mathcal{J} on all individual and function constants, and for all n -ary predicate constants \mathbf{P} , let $(\mathcal{J}'\mathbf{P})d_1 \dots d_n = \sim (\mathcal{J}\mathbf{P})d_1 \dots d_n$ for all $d_1, \dots, d_n \in \mathcal{D}$. Let $\mathbf{M}' = \langle \mathcal{D}, \mathcal{J}' \rangle$. Then for any wff \mathbf{A} and assignment φ , $\mathcal{V}_\varphi^\mathcal{M} \mathbf{A} = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{A}$.

NOTES:

- (1) In accord with the convention introduced in §25, for simplicity we are

assuming that no propositional, predicate, or function variables occur in wffs of \mathcal{F} .

- (2) Clearly φ can be regarded as an assignment into \mathcal{M}' as well as into \mathcal{M} .

Proof by induction on the construction of \mathbf{A} :

Case. \mathbf{A} is an atomic wff $\mathbf{P}t_1 \dots t_n$. It is easy to see by induction that $\mathcal{V}_\varphi^{\mathcal{M}'} t = \mathcal{V}_\varphi^{\mathcal{M}} t$ for each term t . Let

$$d_i = \mathcal{V}_\varphi^{\mathcal{M}} t_i = \mathcal{V}_\varphi^{\mathcal{M}'} t_i \text{ for } i = 1, \dots, n.$$

$$\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A}^d = \mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A} = (\mathcal{J}\mathbf{P})d_1 \dots d_n = \sim (\mathcal{J}'\mathbf{P})d_1 \dots d_n = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{A}.$$

Case. \mathbf{A} is $\sim \mathbf{B}$.

$$\begin{aligned} \mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A}^d &= \mathcal{V}_\varphi^{\mathcal{M}} \sim \mathbf{B}^d = \sim \mathcal{V}_\varphi^{\mathcal{M}} \mathbf{B}^d \\ &= \sim \sim \mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{B} \quad (\text{by inductive hypothesis}) \\ &= \sim \mathcal{V}_\varphi^{\mathcal{M}'} \sim \mathbf{B} = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{A}. \end{aligned}$$

Case. \mathbf{A} is $[\mathbf{B} \vee \mathbf{C}]$.

$$\begin{aligned} \mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A}^d &= \mathcal{V}_\varphi^{\mathcal{M}} [\mathbf{B}^d \wedge \mathbf{C}^d] = \mathcal{V}_\varphi^{\mathcal{M}} \sim [\sim \mathbf{B}^d \vee \sim \mathbf{C}^d] \\ &= \sim (\sim (\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{B}^d) \vee \sim (\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{C}^d)) \\ &= \sim (\sim \sim (\mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{B}) \vee \sim \sim (\mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{C})) \quad (\text{by inductive hypothesis}) \\ &= \sim \mathcal{V}_\varphi^{\mathcal{M}'} [\mathbf{B} \vee \mathbf{C}] = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{A}. \end{aligned}$$

Case. \mathbf{A} is $\forall \mathbf{x} \mathbf{B}$. We must show that $\mathcal{V}_\varphi^{\mathcal{M}} \exists \mathbf{x} \mathbf{B}^d = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \forall \mathbf{x} \mathbf{B}$. Suppose that $\mathcal{V}_\varphi^{\mathcal{M}} \exists \mathbf{x} \mathbf{B}^d = \top$. Then by X2300 there is an assignment ψ which agrees with φ off \mathbf{x} such that $\mathcal{V}_\psi^{\mathcal{M}} \mathbf{B}^d = \top$. By inductive hypothesis, $\mathcal{V}_\psi^{\mathcal{M}} \mathbf{B}^d = \sim \mathcal{V}_\psi^{\mathcal{M}'} \mathbf{B}$, so $\mathcal{V}_\psi^{\mathcal{M}'} \mathbf{B} = \mathbf{F}$. Hence $\mathcal{V}_\varphi^{\mathcal{M}'} \forall \mathbf{x} \mathbf{B} = \mathbf{F}$, so $\sim \mathcal{V}_\varphi^{\mathcal{M}'} \forall \mathbf{x} \mathbf{B} = \top$. For the proof in the other direction, suppose that $\sim \mathcal{V}_\varphi^{\mathcal{M}'} \forall \mathbf{x} \mathbf{B} = \top$. Then $\mathcal{V}_\varphi^{\mathcal{M}'} \forall \mathbf{x} \mathbf{B} = \mathbf{F}$, so there is an assignment ψ which agrees with φ off \mathbf{x} such that $\mathcal{V}_\psi^{\mathcal{M}'} \mathbf{B} = \mathbf{F}$. By inductive hypothesis $\mathcal{V}_\psi^{\mathcal{M}} \mathbf{B}^d = \sim \mathcal{V}_\psi^{\mathcal{M}'} \mathbf{B} = \top$, so $\mathcal{V}_\varphi^{\mathcal{M}} \exists \mathbf{x} \mathbf{B}^d = \top$. Thus $\mathcal{V}_\varphi^{\mathcal{M}} (\forall \mathbf{x} \mathbf{B})^d = \mathcal{V}_\varphi^{\mathcal{M}} \exists \mathbf{x} \mathbf{B}^d = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \forall \mathbf{x} \mathbf{B}$, as required. ■

4003 Corollary. If $\vdash A$, and B is any dual of A , then $\vdash \sim B$.

Proof: Let \mathcal{M} be any interpretation, and define \mathcal{M}' as in 4002. Let φ be any assignment. $\mathcal{V}_\varphi^{\mathcal{M}} \sim A^d = \sim \mathcal{V}_\varphi^{\mathcal{M}} A^d = \sim \sim \mathcal{V}_\varphi^{\mathcal{M}'} A$ (by 4002) = T since $\models A$. Hence $\models \sim A^d$, so $\vdash \sim A^d$ by Gödel's Completeness Theorem (2510). Hence $\vdash \sim B$ by 4000 and Rule P. ■

REMARK. 4003 can also be proved by a purely syntactic argument. See Exercise X4002.

4004 Corollary. If $\vdash A \supset B$ then $\vdash B^d \supset A^d$.

Proof: Suppose $\vdash A \supset B$. Then $\vdash \sim [A^d \not\subset B^d]$ by 4003, so $\vdash B^d \supset A^d$ by Rule P. ■

EXAMPLE: $\vdash \forall x Px \supset Py$ by Axiom 4, so $\vdash Py \supset \exists x Px$ (which is an instance of 2123).

4005 Corollary. If $\vdash A \equiv B$, then $\vdash A^d \equiv B^d$

Proof: Suppose $\vdash A \equiv B$. Then $\vdash \sim [A^d \not\equiv B^d]$ by 4003, so $\vdash A^d \equiv B^d$ by Rule P. ■

EXERCISES

X4000. Complete the proof of Proposition 4000.

X4001. Give a semantic proof of 4001 (using Gödel's Completeness Theorem) by using 4002 to prove that $\mathcal{V}_\varphi^{\mathcal{M}} A^{dd} = \mathcal{V}_\varphi^{\mathcal{M}} A$ for all \mathcal{M} and φ .

X4002. Give a completely syntactic proof of 4003 as follows: Let A_1, \dots, A_n be a proof of A ; show that $\vdash \sim A_i^d$ for $1 \leq i \leq n$ by complete induction on i .

X4003. Give examples to illustrate Corollaries 4003, 4004, and 4005.

X4004. Prove or refute the following conjecture about wffs of the system \mathcal{P} of propositional calculus. For any assignment φ of truth values to propositional variables, let $\varphi' p = \sim \varphi p$ for all p . Suppose A is any wff of \mathcal{P} such that $\mathcal{V}_\varphi A = \sim \mathcal{V}_{\varphi'} A$ for every assignment φ . Then there is a propositional variable q such that $\models A \equiv q$ or $\models A \equiv \sim q$.

§41. Craig's Interpolation Theorem

As in §34, for this section it is convenient to use a formulation of first-order logic which we call \mathcal{F}_f , whose primitive symbols are those of \mathcal{F} plus a 0-ary propositional connective f which is false in all interpretations. (With appropriate minor modifications, the theorem below also applies to formulations of \mathcal{F} which do not contain f .) Note that we regard f as a connective, not a propositional constant. It is readily verified that Smullyan's Unifying Principle (2507) applies to \mathcal{F}_f if one modifies condition (a) in the definition of abstract consistency class to read:

(a') $f \notin \mathcal{S}$, and if A is atomic, then $A \notin \mathcal{S}$ or $\sim A \notin \mathcal{S}$.

We slightly extend the terminology introduced in §21 by defining an occurrence of a constant or variable in a wff A of \mathcal{F}_f to be *positive* [*negative*] in A iff that occurrence is in the scope of an even [odd] number of occurrences of \sim in A .

DEFINITION. A sentence J is an *interpolation sentence* for a sentence $[H \supset K]$ iff the following conditions are all satisfied:

- (1) $\models H \supset J$ and $\models J \supset K$;
- (2) each individual constant which occurs in J occurs in both H and K ;
- (3) each propositional or predicate constant which has a positive [*negative*] occurrence in J also has a positive [*negative*] occurrence in both H and K .

EXAMPLE: J is an interpolation sentence for $[M \wedge J] \supset [J \vee N]$.

REMARK. If $\models K$, then $\models H \supset K$ for any sentence H . Nevertheless, $H \supset K$ has an interpolation sentence, namely $\sim f$. Similarly, if $\models \sim H$, then f is an interpolation sentence for $H \supset K$.

4100 Craig-Lyndon Interpolation Theorem.¹ If $H \supset K$ is a valid sentence of \mathcal{F}_f , then there is an interpolation sentence for $H \supset K$.

Proof: If \mathcal{S}_1 and \mathcal{S}_2 are sets of sentences, we shall say that a sentence J is an interpolation sentence for the ordered pair $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ iff

- (1) $\mathcal{S}_1 \cup \{\sim J\}$ and $\mathcal{S}_2 \cup \{J\}$ are unsatisfiable;

¹[Craig, 1957], [Lyndon, 1959]. Also see [Henkin, 1963a].

- (2) each individual constant which occurs in \mathbf{J} occurs in some sentence of \mathcal{S}_1 and in some sentence of \mathcal{S}_2 ;
- (3) each propositional or predicate constant which has a positive [negative] occurrence in \mathbf{J} also has a positive [negative] occurrence in some sentence of \mathcal{S}_1 and a negative [positive] occurrence in some sentence of \mathcal{S}_2 .

It is readily verified that \mathbf{J} is an interpolation sentence for $\langle \{\mathbf{C}_1, \dots, \mathbf{C}_m\}, \{\sim \mathbf{D}_1, \dots, \sim \mathbf{D}_n\} \rangle$ iff it is an interpolation sentence for $\mathbf{C}_1 \wedge \dots \wedge \mathbf{C}_m \supset \mathbf{D}_1 \vee \dots \vee \mathbf{D}_n$. Also note that :

- (*1) \mathbf{J} is an interpolation sentence for $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ iff $\sim \mathbf{J}$ is an interpolation sentence for $\langle \mathcal{S}_2, \mathcal{S}_1 \rangle$.
- (*2) If \mathbf{J} is an interpolation sentence for $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ and $\mathcal{S}_1 \subseteq \mathcal{U}_1$ and $\mathcal{S}_2 \subseteq \mathcal{U}_2$, then \mathbf{J} is an interpolation sentence for $\langle \mathcal{U}_1, \mathcal{U}_2 \rangle$.

Let $\Gamma = \{\mathcal{S} \mid \mathcal{S} \text{ is a set of sentences such that for some sets } \mathcal{S}_1 \text{ and } \mathcal{S}_2, \mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \text{ and } \langle \mathcal{S}_1, \mathcal{S}_2 \rangle \text{ has no interpolation sentence}\}$. Note that if \mathcal{S} is a set of sentences, then $\mathcal{S} \notin \Gamma$ iff for all sets \mathcal{S}_1 and \mathcal{S}_2 such that $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$, $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ has an interpolation sentence. We shall show that Γ is an abstract consistency class. The theorem will then follow directly, for if $\mathbf{H} \supset \mathbf{K}$ is a valid sentence, then $\{\mathbf{H}, \sim \mathbf{K}\}$ has no model, so $\{\mathbf{H}, \sim \mathbf{K}\} \notin \Gamma$ by Smullyan's Unifying Principle (2507), so $\langle \{\mathbf{H}\}, \{\sim \mathbf{K}\} \rangle$ has an interpolation sentence, so $\mathbf{H} \supset \mathbf{K}$ does too.

Γ is closed under subsets by (*2).

We next check conditions (a'), (b)-(f) in the definition of an abstract consistency class.

(a') f is an interpolation sentence for $\langle \{f\}, \emptyset \rangle$, so $\{f\} \notin \Gamma$ by (*1) and (*2). Also, a sentence \mathbf{A} is an interpolation sentence for $\langle \{\mathbf{A}\}, \{\sim \mathbf{A}\} \rangle$, and f is an interpolation sentence for $\langle \{\mathbf{A}, \sim \mathbf{A}\}, \emptyset \rangle$, so $\{\mathbf{A}, \sim \mathbf{A}\} \notin \Gamma$ by (*1) and (*2). Hence (a') is true since Γ is closed under subsets.

(b) is clear.

We shall check (c)-(f) by the indirect method. In each case (with various choices of the sentence \mathbf{D}) we shall assume (among other things) that $\mathbf{D} \in \mathcal{S}$ and $\mathcal{S} \in \Gamma$. Hence we can choose \mathcal{S}_1 and \mathcal{S}_2 such that $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ and $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ has no interpolation sentence. Since (by (*1)) $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ has an interpolation sentence iff $\langle \mathcal{S}_2, \mathcal{S}_1 \rangle$ does, without loss of generality we may assume that $\mathbf{D} \in \mathcal{S}_1$. In each case we shall reach a contradiction by showing that $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$ must have an interpolation sentence.

(c) Suppose $[A \vee B] \in S$, $S \in \Gamma$, but (c₁) $S \cup \{A\} \notin \Gamma$ and (c₂) $S \cup \{B\} \notin \Gamma$. We are assuming that $[A \vee B] \in S_1$. By (c₁) $\langle S_1 \cup \{A\}, S_2 \rangle$ has an interpolation sentence J, and by (c₂) $\langle S_1 \cup \{B\}, S_2 \rangle$ has an interpolation sentence K. It is easily seen that $[J \vee K]$ is an interpolation sentence for $\langle S_1, S_2 \rangle$.

(d) Suppose $\sim [A \vee B] \in S$, $S \in \Gamma$, but (d₁) $S \cup \{\sim A, \sim B\} \notin \Gamma$. We are assuming that $\sim [A \vee B] \in S_1$. By (d₁) $\langle S_1 \cup \{\sim A, \sim B\}, S_2 \rangle$ has an interpolation sentence J, and J is also an interpolation sentence for $\langle S_1, S_2 \rangle$.

(e) Suppose $\forall x A \in S$ and $S \in \Gamma$, but there is a variable-free term t such that $S \cup \{S_t^x A\} \notin \Gamma$. We are assuming that $\forall x A \in S_1$. There is an interpolation sentence J for $\langle S_1 \cup \{S_t^x A\}, S_2 \rangle$. Let c_1, \dots, c_n (where $n \geq 0$) be the individual constants which occur in t but not in any wff of S_1 , and let y_1, \dots, y_n be distinct individual variables which do not occur in J. Let J' be $S_{y_1 \dots y_n}^{c_1 \dots c_n} J$ and let K be $\forall y_1 \dots \forall y_n J'$. (Thus K is J if $n = 0$.) We assert that K is an interpolation sentence for $\langle S_1, S_2 \rangle$.

$S_1 \cup \{S_t^x A, \sim J\}$ is unsatisfiable, hence inconsistent, so:

$S_1, S_t^x A \vdash J$	by 2501
$S_1 \vdash S_t^x A \supset J$	Deduction Theorem
$S_1 \vdash \forall x A$	since $\forall x A \in S_1$
$S_1 \vdash S_t^x A$	$\forall I$
$S_1 \vdash J$	Rule P
$S_1 \vdash J'$	since c_1, \dots, c_n do not occur in any sentence of S_1
$S_1 \vdash K$	Gen

Thus $S_1 \cup \{\sim K\}$ is inconsistent, hence unsatisfiable. Also, $\vdash K \supset J$ and $S_2 \cup \{J\}$ is unsatisfiable, so $S_2 \cup \{K\}$ is unsatisfiable.

If d is any individual constant in K, then d occurs in J, so d occurs in some sentence of S_2 and in some sentence of $S_1 \cup \{S_t^x A\}$. Suppose d does not occur in any sentence of S_1 . Then (since $\forall x A \in S_1$) d must occur in t, and so is some c_i . However, none of the c_i 's occurs in K, so this is impossible; thus d must indeed occur in some sentence of S_1 . Condition (3) in the definition of an interpolation sentence clearly satisfied, so K is an interpolation sentence for $\langle S_1, S_2 \rangle$.

(f) Suppose $\sim \forall x A \in S$ and $S \in \Gamma$ and c is an individual constant which does not occur in any sentence of S and $S \cup \{\sim S_c^x A\} \notin \Gamma$. Let J be an interpolation sentence for $\langle S_1 \cup \{\sim S_c^x A\}, S_2 \rangle$. We assert that J is also an interpolation sentence for $\langle S_1, S_2 \rangle$. c does not occur in any sentence of S_2 , so c does not occur in J. Hence it is easy to see that every individual

constant in \mathbf{J} occurs in some sentence of \mathcal{S}_1 as well as some sentence of \mathcal{S}_2 . Clearly \mathbf{J} satisfies condition (3).

To complete the proof we simply need to establish that $\mathcal{S}_1 \cup \{\sim \mathbf{J}\}$ is unsatisfiable. Suppose it is satisfiable, and let Δ be the collection of satisfiable sets of sentences of \mathcal{F}_f . By 2505 and 2512, Δ is an abstract consistency class. $\mathcal{S}_1 \cup \{\sim \mathbf{J}\} \in \Delta$ and $\sim \forall x \mathbf{A} \in \mathcal{S}_1 \cup \{\sim \mathbf{J}\}$, so by condition (f) in the definition of an abstract consistency class, $\mathcal{S}_1 \cup \{\sim \mathbf{J}\} \cup \{\sim S_c^x \mathbf{A}\} \in \Delta$, which is to say that this set is satisfiable. This contradicts the fact that \mathbf{J} is an interpolation sentence for $\langle \mathcal{S}_1 \cup \{\sim S_c^x \mathbf{A}\}, \mathcal{S}_2 \rangle$. Thus $\mathcal{S}_1 \cup \{\sim \mathbf{J}\}$ is unsatisfiable.

Hence \mathbf{J} is indeed an interpolation formula for $\langle \mathcal{S}_1, \mathcal{S}_2 \rangle$. ■

§42. Beth's Definability Theorem

DEFINITION. Let \mathbf{H} be a sentence containing a k -ary predicate constant \mathbf{P} , let \mathbf{P}' be a k -ary predicate constant which does not occur in \mathbf{H} , and let \mathbf{H}' be $S_{\mathbf{P}'}^{\mathbf{P}} \mathbf{H}$.

(1) \mathbf{H} defines \mathbf{P} implicitly iff

$$\vdash \mathbf{H} \wedge \mathbf{H}' \supset \forall x_1 \dots \forall x_k \bullet \mathbf{P} x_1 \dots x_k \equiv \mathbf{P}' x_1 \dots x_k.$$

(2) The wff \mathbf{D} is an *explicit definition of \mathbf{P} from \mathbf{H}* iff

- (a) \mathbf{P} does not occur in \mathbf{D} ;
- (b) no variables other than x_1, \dots, x_k occur free in \mathbf{D} ;
- (c) every constant in \mathbf{D} occurs in \mathbf{H} ;
- (d) $\vdash \mathbf{H} \supset \forall x_1 \dots \forall x_k \bullet \mathbf{P} x_1 \dots x_k \equiv \mathbf{D}$.

(3) \mathbf{P} is *explicitly definable from \mathbf{H}* iff there is an explicit definition of \mathbf{P} from \mathbf{H} .

4200 Beth's Definability Theorem [Beth, 1953]. Let \mathbf{H} be a sentence containing a k -ary predicate constant \mathbf{P} . Then \mathbf{H} defines \mathbf{P} implicitly iff \mathbf{P} is explicitly definable from \mathbf{H} .

Proof: If \mathbf{P} is explicitly definable from \mathbf{H} by a wff \mathbf{D} , then

$$\vdash \mathbf{H} \supset \forall x_1 \dots \forall x_k \bullet \mathbf{P} x_1 \dots x_k \equiv \mathbf{D}$$

$\vdash \mathbf{H}' \supset \forall x_1 \dots \forall x_k \bullet \mathbf{P}' x_1 \dots x_k \equiv \mathbf{D}$ by substitution
(since \mathbf{P} does not occur in \mathbf{D}), so it is easily seen that \mathbf{H} defines \mathbf{P} implicitly.
Suppose \mathbf{H} defines \mathbf{P} implicitly. Then

$$\vdash \mathbf{H} \wedge \mathbf{H}' \supset \forall x_1 \dots \forall x_k \bullet \mathbf{P}x_1 \dots x_k \equiv \mathbf{P}'x_1 \dots x_k.$$

By 2515 we may assume that this theorem is provable in a formulation of \mathcal{F} which contains no predicate or function constants other than \mathbf{P}' and the constants which occur in \mathbf{H} . Let $\mathbf{c}_1, \dots, \mathbf{c}_k$ be distinct individual constants of \mathcal{F} which do not occur in \mathbf{H} .

$$\begin{aligned} &\vdash \mathbf{H} \wedge \mathbf{H}' \supset \bullet \mathbf{P}\mathbf{c}_1 \dots \mathbf{c}_k \equiv \mathbf{P}'\mathbf{c}_1 \dots \mathbf{c}_k \\ &\vdash \mathbf{H} \wedge \mathbf{P}\mathbf{c}_1 \dots \mathbf{c}_k \supset \bullet \mathbf{H}' \supset \mathbf{P}'\mathbf{c}_1 \dots \mathbf{c}_k \end{aligned} \quad \text{Rule P}$$

Let \mathbf{B} be an interpolation sentence (see §41) for the sentence above. By making alphabetical changes of bound variables, if necessary, we may assume that the variables x_1, \dots, x_k do not occur in \mathbf{B} . Let \mathbf{D} be $S_{x_1 \dots x_k}^{\mathbf{c}_1 \dots \mathbf{c}_k} \mathbf{B}$. Note that neither \mathbf{P} nor \mathbf{P}' can occur in \mathbf{D} , and every constant in \mathbf{D} must occur in \mathbf{H} .

$$\begin{aligned} &\vdash \mathbf{H} \wedge \mathbf{P}\mathbf{c}_1 \dots \mathbf{c}_k \supset \mathbf{B} && \text{by the definition of } \mathbf{B} \\ &\vdash \mathbf{H} \wedge \mathbf{P}x_1 \dots x_k \supset \mathbf{D} && \text{by substitution} \\ &\vdash \mathbf{B} \supset \bullet \mathbf{H}' \supset \mathbf{P}'\mathbf{c}_1 \dots \mathbf{c}_k && \text{by the definition of } \mathbf{B} \\ &\vdash \mathbf{D} \supset \bullet \mathbf{H}' \supset \mathbf{P}'x_1 \dots x_k && \text{by substitution} \\ &\vdash \mathbf{D} \supset \bullet \mathbf{H} \supset \mathbf{P}x_1 \dots x_k && \text{by substitution} \\ &\vdash \mathbf{H} \supset \bullet \mathbf{P}x_1 \dots x_k \equiv \mathbf{D} && \text{Rule P} \\ &\vdash \mathbf{H} \supset \forall x_1 \dots \forall x_k \bullet \mathbf{P}x_1 \dots x_k \equiv \mathbf{D} && \supset \forall \text{ Rule (2103)} \end{aligned}$$

Thus \mathbf{D} is an explicit definition of \mathbf{P} from \mathbf{H} . ■

REMARK. Note that if \mathbf{P} is (implicitly or explicitly) definable from \mathbf{H} , then \mathbf{P} has an explicit definition \mathbf{D} in which predicates occur only if they have both positive and negative occurrences in \mathbf{H} . This follows from the fact that the interpolation sentence \mathbf{B} constructed in the proof above has this property.

EXERCISE

X4200. Let $A(G, P)$ be the wff

$$\forall x \forall y [Px \supset \bullet Gxx \vee \sim Gxy] \wedge \forall x [Px \vee \exists z Gxz] \wedge \forall x [Gxx \supset Px]$$

(a) Show that $A(G, P)$ defines P implicitly in terms of G . That is, prove $A(G, P_1) \wedge A(G, P_2) \supset \forall w \bullet P_1w \equiv P_2w$.

(b) Illustrate Beth's Theorem on Definability by finding an explicit definition $B(G, w)$ of P in terms of G from $A(G, P)$.

Prove $A(G, P) \supset \forall w \bullet Pw \equiv B(G, w)$.

Chapter 5

Type Theory

§50. Introduction

So far we have been concerned with first-order logic, and its subsystem propositional calculus, which we might regard as zeroth-order logic. We now wish to discuss higher-order logics.

It seems very natural to extend the system \mathcal{F} of first-order logic by permitting quantification on predicate, propositional, and function variables as well as individual variables. One thus obtains the wffs of a system of *second-order logic*. One might then introduce predicate and function variables of higher type to denote relations and functions whose arguments may be relations and functions of individuals as well as individuals. Thus one would be led to a system of *third-order logic*, and if one permitted quantification with respect to these new variables, one would obtain a system of *fourth-order logic*. This process can be continued indefinitely to obtain logics of arbitrarily high order. Of course, after a while one runs out of different types of letters to use for different types of variables, but this problem can be solved by introducing *type symbols* to indicate the types of variables, and using a letter with type symbol α as subscript for a variable of type α .

To make these ideas precise, we shall briefly describe a system \mathcal{F}^ω of ω -order logic which has all finite order logics as subsystems. We shall not discuss \mathcal{F}^ω very extensively, since we prefer to devote most of our attention to an improved formulation of higher-order logic called \mathcal{Q}_0 which we shall soon introduce. Nevertheless, \mathcal{F}^ω provides a convenient starting point for a discussion of higher-order logics. For the sake of simplicity, \mathcal{F}^ω will have no function symbols. It is well known that assertions about functions can be expressed in terms of relations, since to every n -ary function f there

corresponds an $(n + 1)$ -ary relation R such that for all x_1, \dots, x_n , and y , $Rx_1 \dots x_n y$ iff $fx_1 \dots x_n = y$.

A system of logic which includes logics of all finite orders (but no infinite orders) is known as *ω -order logic*, or *finite type theory*. (There are also systems of transfinite type theory, such as that in [Andrews, 1965], but we shall not discuss them here.) Type theory was invented by Bertrand Russell [Russell, 1908]. In [Whitehead and Russell, 1913], which had extraordinary influence on the development of logic in the early twentieth century, he and Alfred North Whitehead demonstrated that various fundamental parts of mathematics could indeed be formalized in this system. Russell was concerned with providing secure foundations for mathematics, and wished to enforce the *vicious-circle principle* that no totality can contain members defined in terms of itself. He therefore originally advocated using what is now called *ramified type theory*,¹ which is still important in various proof-theoretic studies (such as [Feferman, 1964]) of the consistency of restricted portions of mathematics. However, ramified type theory did not prove adequate for formalizing mathematics in general, and we shall devote our attention to a simplified version of it which is known as *simple type theory*.

Additional information about type theory may be obtained from sources such as [Andrews, 2001], [van Benthem and Doets, 1983], [Hatcher, 1968], [Hindley, 1997], [Jacobs, 1999], [Leivant, 1994], [Mendelson, 1987], [Quine, 1963], [Shapiro, 1991], [Wolfram, 1993], and references cited therein.

The *type symbols* of \mathcal{F}^ω and their *orders* are defined inductively as follows:

- (a) ι is a type symbol (denoting the type of individuals) of order 0.
- (b) o is a type symbol (denoting the type of truth values) of order 1.
- (c) If τ_1, \dots, τ_n are type symbols (with $n \geq 1$), then $(\tau_1 \dots \tau_n)$ is a type symbol (denoting the type of n -ary relations with arguments of types τ_1, \dots, τ_n , respectively), and its order is $1 +$ the maximum of the orders of τ_1, \dots, τ_n . (One may think of “ o ” as “()”.)

The primitive symbols of \mathcal{F}^ω are the improper symbols $[,]$, \sim , \vee , and \forall , a denumerable list of variables of each type, and (optionally) constants of certain types. Variables of type o are called *propositional variables*, and variables of type ι are called *individual variables*.

The formation rules of \mathcal{F}^ω are the following:

¹See [Hazen, 1983] or [Church, 1956, p. 347] for more details.

- (a) Every propositional variable or constant is a wff.
- (b) If $u_{\tau_1 \dots \tau_n}, v_{\tau_1}^1 \dots, v_{\tau_n}^n$ are variables or constants of the types indicated by their subscripts, then $u_{\tau_1 \dots \tau_n} v_{\tau_1}^1 \dots v_{\tau_n}^n$ is a wff.
- (c) If A and B are wffs and u is a variable of any type, then $\sim A$, $[A \vee B]$, and $\forall u A$ are wffs.

The *order* of a variable or constant is the order of its type symbol. For any positive integer m , the wffs of $2m$ th-order logic are the wffs in which no variable or constant of order greater than m occurs, and the wffs of $(2m - 1)$ th-order logic are the wffs of $2m$ th-order logic in which no variable of order m is quantified. (Note that first-order logic, as here defined, is just the result of dropping function symbols from the system \mathcal{F} of §20.)

The definitions of connectives and quantifiers in §10 and §20 can be extended to apply to \mathcal{F}^ω . If x_τ and y_τ are variables or constants of the same type τ , $x_\tau = y_\tau$ is defined to be an abbreviation for

$$\forall z_{(\tau)} [z_{(\tau)} x_\tau \supset z_{(\tau)} y_\tau] \quad (*1)$$

where $z_{(\tau)}$ is a variable of type (τ) . An alternative possibility is to define $x_\tau = y_\tau$ as

$$\forall p_{(\tau\tau)} [\forall z_\tau p_{(\tau\tau)} z_\tau z_\tau \supset p_{(\tau\tau)} x_\tau y_\tau]. \quad (*2)$$

(See Exercise X5007.) Note that the two basic properties of equality which we took as axioms in §26 are reflexivity and substitutivity. *1 defines equality in terms of substitutivity, and *2 defines equality in terms of reflexivity (indeed, as the intersection of all reflexive relations). Thus, all the properties of equality follow from either of the basic properties of substitutivity or reflexivity when these are expressed in a sufficiently powerful way using the expressiveness of higher-order logic.

The rules of inference of \mathcal{F}^ω are:

- (1) **Modus Ponens.** From A and $A \supset B$ to infer B .
- (2) **Generalization.** From A to infer $\forall x A$, where x is a variable of any type.

The axiom schemata of \mathcal{F}^ω are:

- (1) $A \vee A \supset A$
- (2) $A \supset \bullet B \vee A$
- (3) $A \supset B \supset \bullet C \vee A \supset B \vee C$

(4) $\forall x_\tau A \supset S_{y_\tau}^x A$, where y_τ is a variable or constant of the same type as the variable x_τ , and y_τ is free for x_τ in A .

(5) $\forall x[A \vee B] \supset \neg A \vee \forall x B$, where x is any variable not free in A .

(6) **(Comprehension Axioms)**

$\exists u_o [u_o \equiv A]$, where u_o does not occur free in A .

$\exists u_{(\tau_1 \dots \tau_n)} \forall v_{\tau_1}^1 \dots \forall v_{\tau_n}^n [u_{(\tau_1 \dots \tau_n)} v_{\tau_1}^1 \dots v_{\tau_n}^n \equiv A]$, where $u_{(\tau_1 \dots \tau_n)}$ does not occur free in A , and $v_{\tau_1}^1, \dots, v_{\tau_n}^n$ are distinct variables.

(7) **(Axioms of Extensionality)**

$[x_o \equiv y_o] \supset \neg x_o = y_o$

$\forall w_{\tau_1}^1 \dots \forall w_{\tau_n}^n [x_{(\tau_1 \dots \tau_n)} w_{\tau_1}^1 \dots w_{\tau_n}^n \equiv y_{(\tau_1 \dots \tau_n)} w_{\tau_1}^1 \dots w_{\tau_n}^n] \supset \neg x_{(\tau_1 \dots \tau_n)} = y_{(\tau_1 \dots \tau_n)}$

The axioms of k th-order logic are those axioms of \mathcal{F}^ω which are wffs of k th-order logic. Note that Comprehension Axioms first appear in second-order logic, and Axioms of Extensionality first appear in fourth-order logic.

Note the principal features which are added to first-order logic in order to obtain this formulation of higher-order logic:

1. Variables of arbitrarily high orders.
2. Quantification on variables of all types.
3. Comprehension Axioms.
4. Axioms of Extensionality.

One may also wish to assume additional axioms, such as the Axiom Schema of Choice, an Axiom of Infinity, and perhaps even the Continuum Hypothesis, though there is room for argument whether these are axioms of pure logic or of mathematics.

It should be remarked that one can restrict the formation rules of \mathcal{F}^ω by requiring that $n = 1$ in the definitions above, and obtain a still simpler system of ω -order logic which is nevertheless adequate to express most mathematical ideas, since one can define ordered pairs (and, more generally, n -tuples) in various ways, and then represent an n -ary relation B by the monadic relation P such that $Bx^1 \dots x^n$ iff $P(x_1, \dots, x_n)$. Thus, $P(x^1, \dots, x^n)$ would be an abbreviation for $\exists u [Pu \wedge u = \langle x^1, \dots, x^n \rangle]$, where $u = \langle x^1, \dots, x^n \rangle$ is defined in some appropriate way, and suitable types are assigned to u and P .

If one wishes to thus restrict the language so that all predicates are monadic, one may find it congenial to write $v_\tau \in u_{(\tau)}$ rather than $u_{(\tau)}v_\tau$,

and to say “ v_τ is in the set $u_{(\tau)}$ ” rather than “ v_τ has the property $u_{(\tau)}$ ”. Of course, this is a mere notational change. A more radical change is obtained by also dropping the type symbols, so that all variables are of the same type. (Propositional variables are customarily dispensed with.) One thus obtains Naive Axiomatic Set Theory, which can be formalized in first-order logic. \in is regarded as a binary predicate constant, and all atomic wffs are of the form $x \in y$, where x and y are (individual) variables. $x \in y$ is interpreted to mean “ x is a member of the set y ”, which is of course false if y is not a set at all. The Comprehension Axiom Schema now takes the simple form $\exists u \forall v [v \in u \equiv A]$, where u is not free in the wff A . This axiom schema asserts that for any wff A in which v occurs free, there is a set u consisting of all those elements v of which A is true. Unfortunately, the axiom schema guarantees the existence of too many sets. In particular, $\exists u \forall v [v \in u \equiv \sim v \in v]$, i.e., there is a set whose members are all those elements which are not members of themselves. Choosing such a set u , we conclude that $u \in u \equiv \sim u \in u$, which is a contradiction, so Naive Axiomatic Set Theory is inconsistent! The paradox above was discovered by Bertrand Russell, and is known as Russell’s Paradox.

One can restrict the Comprehension Axiom Schema, and take only certain carefully chosen instances of it as axioms, to obtain apparently consistent formulations of Axiomatic Set Theory. Alternatively, one can keep type symbols, so that Russell’s paradox (and various other paradoxes) cannot be expressed at all within the formal system. Each approach leads to systems which are adequate for formalizing mathematics, and each approach has advantages for certain purposes. At present, Axiomatic Set Theory is more popular with mathematicians simply because it is much better known. However, most mathematicians do make mental distinctions between different types of mathematical objects, and use different types of letters to denote them, so it might be claimed that type theory provides a more natural formalization of mathematics as it is actually practiced. In addition, explicit syntactic distinctions between expressions denoting intuitively different types of mathematical entities are very useful in computerized systems for exploring and applying mathematics (on a theoretical rather than purely computational level). Of course, type symbols need not be explicitly written as long as one knows the types of the variables and constants.

Therefore we shall turn our attention to finding a formulation of type theory which is as expressive as possible, allowing mathematical ideas to be expressed precisely with a minimum of circumlocutions, and which is as simple and economical as is possible without sacrificing expressiveness. The reader will observe that the formal language we find arises very naturally

from a few fundamental design decisions.

Since function symbols are very useful in practice, we shall admit them into the language. Indeed, for all types α and β , we shall have a type $(\alpha\beta)$ of functions from elements of type β to elements of type α .² (Of course, we are now giving the type symbol $(\alpha\beta)$ a different meaning from that used for the system \mathcal{F}^ω .)

Since the values of functions can themselves be functions, by using a device of Schönfinkel [Schönfinkel, 1924] we can avoid explicitly admitting into our language functions of more than one argument. For example, if f is a function of two arguments, for each element x of the left domain of f there is a function g (depending on x) such that for each element y of the right domain of f , $gy = fxy$. We may now write $g = fx$, and regard f as a function of a single argument, whose value for any argument x in its domain is a function fx , whose value for any argument y in its domain is fxy .

For a more explicit example, consider the function $+$ which carries any pair of natural numbers to their sum. We may denote this function by $+((\sigma\sigma)\sigma)$, where σ is the type of natural numbers. Given any number x , $[+((\sigma\sigma)\sigma)x]$ is the function which, when applied to any number y , gives the value $[[+((\sigma\sigma)\sigma)x]y]$, which is ordinarily abbreviated as $x + y$. Thus $[+((\sigma\sigma)\sigma)x]$ is the function of one argument x which adds x to any number. When we think of $+((\sigma\sigma)\sigma)$ as a function of one argument, we see that it maps any number x to the function $[+((\sigma\sigma)\sigma)x]$.

More generally, if f is a function which maps n -tuples $\langle x_{\beta_1}^1, \dots, x_{\beta_n}^n \rangle$ of elements of types β_1, \dots, β_n , respectively, to elements of type α , we may assign to f the type $((\dots(\alpha\beta_n) \dots \beta_2)\beta_1)$. (We may omit parentheses, using the convention of association to the left, and write this type symbol simply as $(\alpha\beta_n \dots \beta_1)$.)

A set or property can be represented by a function which maps elements to truth values, so that an element is in the set, or has the property, in question iff the function representing the set or property maps that element to truth. We take o as the type symbol denoting the type of truth values, so we may speak of any function of type $(o\alpha)$ as a *set* of elements of type α . A function of type $((o\alpha)\beta)$ is a binary relation between elements of type β and elements of type α . For example, if σ is the type of the natural numbers, and $<$ is the order relation between natural numbers, $<$ has type $(o\sigma\sigma)$, and for all natural numbers x and y , $<xy$ (which we ordinarily write as $x < y$)

²Some authors write $(\beta \rightarrow \alpha)$ instead of $(\alpha\beta)$. It is easy to remember the meaning of the notation $(\beta \rightarrow \alpha)$, but when it is used in long wffs with many variables, the numerous arrows provide useless visual clutter.

has the value truth iff x is less than y . Of course, $<$ can also be regarded as the function which maps each natural number x to the set $\{y \mid x < y\}$ of all natural numbers y such that x is less than y . Thus we shall regard sets, properties and relations as particular kinds of functions.

In first-order logic we let *terms* denote individuals, while *wffs* expressed statements which took on truth values, i.e., for given interpretations and assignments, wffs denoted truth values. Since we now have many more types under consideration, we need a more uniform terminology, so we shall call expressions which denote elements of type α *wffs of type α* . Thus, statements of type theory will be wffs of type o .

The Comprehension Axiom Schema says that for any wff A_o of type o in which $u_{o\beta}$ does not occur free, $\exists u_{o\beta} \forall v_\beta [u_{o\beta} v_\beta \equiv A_o]$. We may generalize this to assert the existence of functions as well as sets, and state it as follows: $\exists u_{\alpha\beta} \forall v_\beta [u_{\alpha\beta} v_\beta = A_\alpha]$ for any wff A_α of type α in which $u_{\alpha\beta}$ is not free. We shall denote the function $u_{\alpha\beta}$ whose existence is thus asserted by $[\lambda v_\beta A_\alpha]$. Thus λv_β is a new variable-binder, like $\forall v_\beta$ or $\exists v_\beta$ (but with a quite different meaning, of course); λ is known as an *abstraction operator*. $[\lambda v_\beta A_\alpha]$ denotes the function whose value on any argument v_β is A_α , where v_β may occur free in A_α . For example, $[\lambda n_\sigma . n_\sigma^2 + 3]$ denotes the function whose value on any natural number n is $n^2 + 3$. Hence when we apply this function to the number 5 we obtain $[\lambda n_\sigma . n_\sigma^2 + 3]5 = 5^2 + 3 = 28$.

When written in terms of λ -notation, the Comprehension Axiom becomes $\forall v_\beta [\lambda v_\beta A_\alpha] v_\beta = A_\alpha$. From this by universal instantiation we obtain $[\lambda v_\beta A_\alpha] B_\beta = S_{B_\beta}^{v_\beta} A_\alpha$ (provided B_β is free for v_β in A_α), i.e., the result of applying the function $[\lambda v_\beta A_\alpha]$ to the argument B_β is $S_{B_\beta}^{v_\beta} A_\alpha$. The process of replacing $[\lambda v_\beta A_\alpha] B_\beta$ by $S_{B_\beta}^{v_\beta} A_\alpha$ (or vice-versa) is known as λ -conversion. Of course, when A_o is a wff of type o , $[\lambda v_\beta A_o]$ denotes the set of all elements v_β (of type β) of which A_o is true; this set may also be denoted by $\{v_\beta \mid A_o\}$. In familiar set-theoretic notation, $[\lambda v_\beta A_o] B_\beta = S_{B_\beta}^{v_\beta} A_o$ would be written $B_\beta \in \{v_\beta \mid A_o\} \equiv S_{B_\beta}^{v_\beta} A_o$. (By the Axiom of Extensionality for truth values, when C_o and D_o are of type o , $C_o \equiv D_o$ is equivalent to $C_o = D_o$.)

It is interesting to note that the quantifiers and connectives need no longer be regarded as improper symbols, but can be assigned types and can be denoted by constants of these types. The function n_{oo} whose existence is asserted by the comprehension axiom $\exists n_{oo} \forall p_o [n_{oo} p_o \equiv \sim p_o]$, and which is denoted by $[\lambda p_o . \sim p_o]$, is clearly just the negation function, so we see that the negation function has type (oo) . This is natural, since it maps truth values to truth values. Similarly, disjunction and conjunction (etc.) are

binary functions from truth values to truth values, so they have type (ooo) .

The statement $\forall x_\alpha A_o$ is true iff the set $[\lambda x_\alpha A_o]$ contains all elements of type α . The comprehension axiom $\exists p_{o(o\alpha)} \forall s_{(o\alpha)} [p_{o(o\alpha)} s_{o\alpha} \equiv \forall x_\alpha \cdot s_{o\alpha} x_\alpha]$ asserts the existence of a property $\Pi_{o(o\alpha)}$ of sets such that $\forall s_{o\alpha} [\Pi_{o(o\alpha)} s_{o\alpha} \equiv \forall x_\alpha \cdot s_{o\alpha} x_\alpha]$, i.e., a set $s_{o\alpha}$ has the property $\Pi_{o(o\alpha)}$ iff $s_{o\alpha}$ contains all elements of type α . In particular, $\Pi_{o(o\alpha)} [\lambda x_\alpha A_o] \equiv \forall x_\alpha [[\lambda x_\alpha A_o] x_\alpha]$. Since $[\lambda x_\alpha A_o] x_\alpha \equiv A_o$ by λ -conversion, we have $\Pi_{o(o\alpha)} [\lambda x_\alpha A_o] \equiv \forall x_\alpha A_o$, so if we introduce $\Pi_{o(o\alpha)}$ as a constant, we can define $\forall x_\alpha$ in terms of it, and λ is the only variable-binder that we need.

A formulation of type theory based on these ideas was introduced by Church in [Church, 1940], and proved complete by Henkin in [Henkin 1950].³ In this system equality can be defined using connectives and quantifiers. However, it is also possible to define connectives and quantifiers in terms of equality. Equality is a very basic and simple notion, so instead of using Church's formulation of type theory, we shall use a slight variant of it (first introduced in [Henkin, 1963b] and simplified in [Andrews, 1963]) in which equality is taken as the basic primitive notion. (The last paragraph of [Andrews, 1972a] gives an additional reason for this choice of primitive symbols.)

We shall denote the binary relation of equality between elements of type α by the constant $Q_{o\alpha\alpha}$. (Naturally, elements which are the same have the same type.) We can easily find an expression T_o to denote truth; any expression of the form $\mathbf{A} = \mathbf{A}$ will do. Then $[\lambda x_\alpha T_o]$ denotes the set of all elements of type α , so we may regard $\forall x_\alpha A_o$ as an abbreviation for $[\lambda x_\alpha T_o] = [\lambda x_\alpha A_o]$. If anything is false, the assertion $\forall x_o x_o$ that everything is true must be false, so we let this wff, which we also write as F_o , denote falsehood. The expression $[\lambda g_{ooo} \cdot g_{ooo} x_o y_o]$ can be used to represent the ordered pair $\langle x_o, y_o \rangle$, and the conjunction $x_o \wedge y_o$ is true iff x_o and y_o are both true, i.e., iff $\langle T_o, T_o \rangle = \langle x_o, y_o \rangle$. Hence $x_o \wedge y_o$ can be expressed by the formula $[\lambda g_{ooo} \cdot g_{ooo} T_o T_o] = [\lambda g_{ooo} \cdot g_{ooo} x_o y_o]$. From falsehood, conjunction, and equivalence (the equality relation Q_{ooo} between truth values), all other propositional connectives can be defined. Of course, primitive constants could be introduced to denote all these concepts, but the resulting system would be redundant.

There is one additional notation which we wish to have available in our language. If \mathbf{A}_o is a wff of type o in which the variable y_i occurs free, and there is a unique element y_i of which \mathbf{A}_o is true, we denote that element by

³See exercise X8030.

$[\gamma y, \mathbf{A}_o]$, and call it “the y_i such that \mathbf{A}_o ”.⁴ What we have here, in essence, is a function which maps the set $[\lambda y, \mathbf{A}_o]$ to its unique element. We shall denote such a function by $\iota_{i(o)}$ and define $[\gamma y, \mathbf{A}_o]$ to be an abbreviation for $\iota_{i(o)}[\lambda y, \mathbf{A}_o]$. $\iota_{i(o)}$ is called a *description operator* (for individuals), since \mathbf{A}_o describes $[\gamma y, \mathbf{A}_o]$. (We shall see in §53 how description operators for other types can be introduced by definition.) Actually, if there is a unique individual y_i in the set $[\lambda y, \mathbf{A}_o]$, then this set is simply the set whose only member is y_i , which can be denoted by $Q_{on}y_i$, since for any individual x_i , x_i is in $Q_{on}y_i$ iff $y_i = x_i$, i.e., iff $Q_{on}y_i x_i$. Thus the assertion that $\iota_{i(o)}$ maps one-element sets to their unique members can be expressed by the simple proposition that $\iota_{i(o)}[Q_{on}y_i] = y_i$ (for each individual y_i). (Thus $\iota_{i(o)}$ is an inverse for Q_{on} . Q_{on} maps individuals to the unit sets containing them, and $\iota_{i(o)}$ maps these sets back to their unique members.)

In §51 we shall present a system \mathcal{Q}_0 of type theory based on these ideas.

EXERCISES

Give proofs for the theorems \mathcal{F}^ω stated below. Use appropriate analogues of the derived rules of inference of §21.

X5000. $\vdash x_\tau = x_\tau$

X5001. $\vdash x_\tau = y_\tau \supset \bullet S_{x_\tau}^{\mathbf{z}_\tau} \mathbf{A} \supset S_{y_\tau}^{\mathbf{z}_\tau} \mathbf{A}$ provided x_τ and y_τ are free for \mathbf{z}_τ in \mathbf{A} .

X5002. $\vdash x_\tau = y_\tau \supset y_\tau = x_\tau$

X5003. $\vdash x_\tau = y_\tau \supset \bullet y_\tau = z_\tau \supset x_\tau = z_\tau$

X5004. $\vdash x_{(\tau)} = y_{(\tau)} \supset \forall z_\tau \bullet xz \supset yz$

X5005. $\vdash x_{(\tau)} = y_{(\tau)} \supset \forall z_\tau \bullet yz \supset xz$

X5006. $\vdash x_{(\tau)} = y_{(\tau)} \equiv \forall z_\tau \bullet xz \equiv yz$

X5007. $\vdash x_\tau = y_\tau \equiv \forall p_{(\tau\tau)} [\forall z_\tau p_{(\tau\tau)} z_\tau z_\tau \supset p_{(\tau\tau)} x_\tau y_\tau]$

One formulation of the Axiom of Choice says that for each collection u of nonempty pairwise disjoint sets (whose elements have type τ), there is a set r containing exactly one element from each set in u . This can be expressed in \mathcal{F}^ω by

⁴ A strong tradition of using the inverted iota for this purpose was established by its use in [Whitehead and Russell, 1913].

$$\begin{aligned}
 AC_1^\tau & : \forall u_{((\tau))} \bullet \forall p_{(\tau)} [u_{((\tau))} p_{(\tau)} \supset \exists x_\tau p_{(\tau)} x_\tau] \\
 & \wedge \forall p_{(\tau)} \forall q_{(\tau)} [u_{((\tau))} p_{(\tau)} \wedge u_{((\tau))} q_{(\tau)} \wedge p_{(\tau)} \neq q_{(\tau)}] \\
 & \supset \forall x_\tau \sim \bullet p_{(\tau)} x_\tau \wedge q_{(\tau)} x_\tau] \\
 & \supset \exists r_{(\tau)} \forall p_{(\tau)} \bullet u_{((\tau))} p_{(\tau)} \supset \exists_1 x_\tau \bullet p_{(\tau)} x_\tau \wedge r_{(\tau)} x_\tau.
 \end{aligned}$$

A second formulation of the Axiom of Choice says that there is a function (a special sort of binary relation) f which chooses from each nonempty set p (whose elements have type τ) a member of that set. This can be expressed in \mathcal{F}^ω by

$$\begin{aligned}
 AC_2^\tau & : \exists f_{((\tau)\tau)} \forall p_{(\tau)} \bullet [\exists x_\tau p_{(\tau)} x_\tau \supset \exists_1 x_\tau f_{((\tau)\tau)} p_{(\tau)} x_\tau] \\
 & \wedge \forall x_\tau \bullet f_{((\tau)\tau)} p_{(\tau)} x_\tau \supset p_{(\tau)} x_\tau
 \end{aligned}$$

X5008. $\vdash AC_2^\tau \supset AC_1^\tau$

X5009. $\vdash AC_1^{((\tau)\tau)} \supset AC_2^\tau$

§51. The Primitive Basis of \mathcal{Q}_0

We now describe a system \mathcal{Q}_0 of type theory which is based on the ideas discussed in §50. We shall intersperse with our formal description of the language \mathcal{Q}_0 occasional informal comments concerning the intended interpretations of the language.

We use α , β , γ , etc., as syntactical variables ranging over *type symbols*, which are defined inductively as follows:

- (a) i is a type symbol (denoting the type of individuals.)
- (b) o is a type symbol (denoting the type of truth values).
- (c) If α and β are type symbols, then $(\alpha\beta)$ is a type symbol (denoting the type of functions from elements of type β to elements of type α).

The *primitive symbols* of \mathcal{Q}_0 are the following:

- (a) Improper symbols: [] λ
- (b) For each type symbol α , a denumerable list of *variables* of type α :
 $f_\alpha \ g_\alpha \ h_\alpha \dots x_\alpha \ y_\alpha \ z_\alpha \ f_\alpha^1 \ g_\alpha^1 \dots z_\alpha^1 \ f_\alpha^2 \dots$

We shall use $f_\alpha, g_\alpha, \dots, x_\alpha, y_\alpha, z_\alpha$, etc., as syntactical variables for variables of type α .

- (c) Logical constants: $Q_{((\alpha\alpha)\alpha)} \ i_{(i(\alpha))}$

The types of these constants are indicated by their subscripts. $Q_{((\alpha\alpha)\alpha)}$ denotes the identity relation between elements of type α . $i_{(i(\alpha))}$ denotes a description or selection operator for individuals.

- (d) In addition, there may be other constants of various types, which will be called *nonlogical constants* or *parameters*. Each choice of parameters determines a particular formulation of \mathcal{Q}_0 .

We write wff_α as an abbreviation for *wff of type α* , and use $A_\alpha, B_\alpha, C_\alpha$, etc., as syntactical variables ranging over $wffs_\alpha$, which are defined inductively as follows:

- (a) A primitive variable or constant of type α is a wff_α .
- (b) $[A_{\alpha\beta} B_\beta]$ is a wff_α .
- (c) $[\lambda x_\beta A_\alpha]$ is a $wff_{(\alpha\beta)}$.

In the intended interpretations of \mathcal{Q}_0 , $[A_{\alpha\beta} B_\beta]$ denotes the value of the function denoted by $A_{\alpha\beta}$ at the argument denoted by B_β . Also, $[\lambda x_\beta A_\alpha]$ denotes the function whose value at any element x_β of type β is (the denotation of) A_α (which depends on x_β , in general). More precise discussion of interpretations of \mathcal{Q}_0 will occur in §54.

5100 Principle of Induction on the Construction of a Wff. Let \mathcal{R} be a property of formulas, and let $\mathcal{R}(A)$ mean that A has property \mathcal{R} . Suppose

- (1) Every formula consisting of a single variable or constant standing alone has property \mathcal{R} .
- (2) Whenever $A_{\alpha\beta}$ is a $wff_{(\alpha\beta)}$ and B_β is a wff_β and $\mathcal{R}(A_{\alpha\beta})$ and $\mathcal{R}(B_\beta)$, then $\mathcal{R}([A_{\alpha\beta} B_\beta])$.
- (3) Whenever x_β is a variable and $\mathcal{R}(A_\alpha)$, then $\mathcal{R}([\lambda x_\beta A_\alpha])$.

Then every wff of \mathcal{Q}_0 has property \mathcal{R} .

Proof: Left to the reader. ■

Brackets, parentheses in type symbols, and type symbols may be omitted when no ambiguity is thereby introduced. The type symbol o shall usually be omitted, and we often use \mathbf{A} , \mathbf{B} , \mathbf{C} , etc., as syntactical variables for wffs _{o} unless otherwise stated. A dot in an abbreviation for a wff stands for a left bracket whose mate is as far to the right as is consistent with the pairing of brackets already present and with the formula being well-formed. Otherwise brackets and parentheses are to be restored using the convention of association to the left.

We introduce the following definitions and abbreviations:

$[\mathbf{A}_\alpha = \mathbf{B}_\alpha]$	stands for $[Q_{o\alpha\alpha} \mathbf{A}_\alpha \mathbf{B}_\alpha]$.
$[\mathbf{A} \equiv \mathbf{B}]$	stands for $[Q_{ooo} \mathbf{A} \mathbf{B}]$.
T_o	stands for $[Q_{ooo} = Q_{ooo}]$.
F_o	stands for $[\lambda x_o T] = [\lambda x_o x_o]$.
$\Pi_{o(o\alpha)}$	stands for $[Q_{o(o\alpha)(o\alpha)} [\lambda x_\alpha T]]$.
$[\forall x_\alpha \mathbf{A}]$	stands for $[\Pi_{o(o\alpha)} [\lambda x_\alpha \mathbf{A}]]$.
\wedge_{ooo}	stands for $[\lambda x_o \lambda y_o \bullet [\lambda g_{ooo} \bullet g_{ooo} TT] = [\lambda g_{ooo} \bullet g_{ooo} x_o y_o]]$.
$[\mathbf{A} \wedge \mathbf{B}]$	stands for $[\wedge_{ooo} \mathbf{A} \mathbf{B}]$.
\supset_{ooo}	stands for $[\lambda x_o \lambda y_o \bullet x_o = \bullet x_o \wedge y_o]$.
$[\mathbf{A} \supset \mathbf{B}]$	stands for $[\supset_{ooo} \mathbf{A} \mathbf{B}]$.
\sim_{oo}	stands for $[Q_{ooo} F]$.
\vee_{ooo}	stands for $[\lambda x_o \lambda y_o \bullet \sim \bullet [\sim x_o] \wedge [\sim y_o]]$.
$[\mathbf{A} \vee \mathbf{B}]$	stands for $[\vee_{ooo} \mathbf{A} \mathbf{B}]$.
$[\exists x_\alpha \mathbf{A}]$	stands for $[\sim \forall x_\alpha \sim \mathbf{A}]$.
$[\mathbf{A}_\alpha \neq \mathbf{B}_\alpha]$	stands for $[\sim \bullet \mathbf{A}_\alpha = \mathbf{B}_\alpha]$.

We supplement our conventions for restoring brackets in abbreviations of wffs _{o} so as to incorporate the conventions used previously for propositional calculus and first-order logic. Naturally, brackets must be restored at each stage so that the pairing of brackets already present is not altered, and so that the entire formula is well-formed. After restoring left brackets indicated by dots (and the mates of these brackets), enclose parts of the forms $[\sim \mathbf{A}]$, $[\forall x_\alpha \mathbf{A}]$ and $[\exists x_\alpha \mathbf{A}]$ in such a way that \sim , \forall , and \exists have the smallest possible scope. Then brackets enclosing parts of the form $[\mathbf{A} \wedge \mathbf{B}]$ are to be restored in such a way that each \wedge has the smallest possible scope, starting from the left, so that the law of association to the left governs the relationship of brackets between several such parts. Then brackets enclosing parts of the form $[\mathbf{A} \vee \mathbf{B}]$, followed by parts of the form $[\mathbf{A} \supset \mathbf{B}]$, and finally $[\mathbf{A} \equiv \mathbf{B}]$ (or $[\mathbf{A} = \mathbf{B}]$, where \mathbf{A} and \mathbf{B} are wffs _{o}), are to be similarly restored.

We next state the axioms for Q_0 . A remark about Axiom Schemata 4₁

– 4₅ will be made after some necessary terminology has been introduced.

Axioms for \mathcal{Q}_0

- (1) $g_{oo}T_o \wedge g_{oo}F_o = \forall x_o \bullet g_{oo}x_o$
- (2 α) $[x_\alpha = y_\alpha] \supset \bullet h_{o\alpha}x_\alpha = h_{o\alpha}y_\alpha$
- (3 $\alpha\beta$) $f_{\alpha\beta} = g_{\alpha\beta} = \forall x_\beta \bullet f_{\alpha\beta}x_\beta = g_{\alpha\beta}x_\beta$
- (4₁) $[\lambda x_\alpha B_\beta] A_\alpha = B_\beta$ where B_β is a primitive constant or variable distinct from x_α
- (4₂) $[\lambda x_\alpha x_\alpha] A_\alpha = A_\alpha$
- (4₃) $[\lambda x_\alpha \bullet B_{\beta\gamma} C_\gamma] A_\alpha = [[\lambda x_\alpha B_{\beta\gamma}] A_\alpha][[\lambda x_\alpha C_\gamma] A_\alpha]$
- (4₄) $[\lambda x_\alpha \bullet \lambda y_\gamma B_\delta] A_\alpha = [\lambda y_\gamma \bullet [\lambda x_\alpha B_\delta] A_\alpha]$ where y_γ is distinct from x_α and from all variables in A_α .
- (4₅) $[\lambda x_\alpha \bullet \lambda x_\alpha B_\delta] A_\alpha = [\lambda x_\alpha B_\delta]$
- (5) $\iota_{i(o)}[Q_{on}y_i] = y_i$

\mathcal{Q}_0 has a single rule of inference, which is the following:

Rule R. From C and $A_\alpha = B_\alpha$ to infer the result of replacing one occurrence of A_α in C by an occurrence of B_α , provided that the occurrence of A_α in C is not (an occurrence of a variable) immediately preceded by λ .

As in other logistic systems, a *proof* of a wff A in \mathcal{Q}_0 is a finite sequence of wffs, the last of which is A , such that every wff in the sequence is an axiom or is inferred from previous wffs by the rule of inference. A *theorem* of \mathcal{Q}_0 is a wff which has a proof in \mathcal{Q}_0 .

An occurrence of x_α is *bound* [*free*] in B_β iff it is [is not] in a wf part of B_β of the form $[\lambda x_\alpha C_\delta]$. A_α is *free for* x_α in B_β iff no free occurrence of x_α in B_β is in a wf part of B_β of the form $[\lambda y_\gamma C_\delta]$ such that y_γ is a free variable of A_α .

A *closed wff* (*cwff*) is a wff without free variables. A *sentence* is a cwff_o.

If $x_{\alpha_1}^1, \dots, x_{\alpha_n}^n$ are distinct variables, $S_{A_{\alpha_1}^1 \dots A_{\alpha_n}^n}^{x_{\alpha_1}^1 \dots x_{\alpha_n}^n} B_\beta$ denotes the result of simultaneously substituting $A_{\alpha_i}^i$ for all free occurrences of $x_{\alpha_i}^i$ in B_β for $1 \leq i \leq n$.

REMARK. Axiom Schemata 4₁ - 4₅ express the basic properties of λ -notation, and will be used in conjunction with Axiom Schema 3 in §52 to derive Theorem Schema 5207, which says that $[\lambda x_\alpha B_\beta] A_\alpha = S_{A_\alpha}^{x_\alpha} B_\beta$, provided that

\mathbf{A}_α is free for x_α in \mathbf{B}_β . 5207 could be taken as an axiom schema in place of 4₁ - 4₅, and for some purposes this would be desirable, since 5207 has a conceptual simplicity and unity which is not apparent in 4₁ - 4₅. However, there is a sense in which it is simpler to state 4₁ - 4₅, which do not require prior understanding of substitution and related notions. This simplicity will facilitate our work in §70, where we shall arithmetize the syntax of \mathcal{Q}_0 .

When one considers the primitive symbols of \mathcal{Q}_0 , it is apparent that from a conceptual point of view, the axioms of \mathcal{Q}_0 are really quite simple and natural. Axiom 1 expresses the idea that Truth and Falsehood are the only truth values, Axioms 2 ^{α} express one of the basic properties of equality, Axioms 3 ^{$\alpha\beta$} are Axioms of Extensionality, Axioms 4₁ - 4₅ express the properties of λ -notation, and Axiom 5 is the Axiom of Descriptions.

If \mathcal{H} is a finite set of wffs_o, we shall write $\mathcal{H} \vdash \mathbf{A}$ as an abbreviation for the phrase “ \mathbf{A} is derivable from the set \mathcal{H} of hypotheses”, and we shall define this concept so that:

- (a) If $\mathbf{A} \in \mathcal{H}$, then $\mathcal{H} \vdash \mathbf{A}$
- (b) If \mathbf{A} is a theorem of \mathcal{Q}_0 , then $\mathcal{H} \vdash \mathbf{A}$.
- (c) (Rule R') If $\mathcal{H} \vdash \mathbf{A}_\alpha = \mathbf{B}_\alpha$, and $\mathcal{H} \vdash \mathbf{C}$, and \mathbf{D} is obtained from \mathbf{C} by replacing one occurrence of \mathbf{A}_α in \mathbf{C} by an occurrence of \mathbf{B}_α , then $\mathcal{H} \vdash \mathbf{D}$, provided that the occurrence of \mathbf{A}_α in \mathbf{C} is not an occurrence of a variable immediately preceded by λ , and the occurrence of \mathbf{A}_α in \mathbf{C} is not in a wf part $[\lambda x_\beta E_\gamma]$ of \mathbf{C} , where x_β is free in a member of \mathcal{H} and free in $[\mathbf{A}_\alpha = \mathbf{B}_\alpha]$.

DEFINITION. A *proof of \mathbf{A} from the hypotheses \mathcal{H}* consists of two finite sequences \mathcal{S}_1 and \mathcal{S}_2 of wffs_o, such that \mathcal{S}_1 is a proof in \mathcal{Q}_0 , and \mathbf{A} is the last member of the sequence \mathcal{S}_2 , and each wff \mathbf{D}^i in \mathcal{S}_2 satisfies at least one of the following conditions:

- (a) \mathbf{D}^i is in \mathcal{H} .
- (b) \mathbf{D}^i is in \mathcal{S}_1 .
- (c) \mathbf{D}^i is obtained by Rule R' from wffs \mathbf{D}^j and \mathbf{D}^k which precede \mathbf{D}^i in the sequence \mathcal{S}_2 .

We cannot replace condition (b) in this definition by the requirement that \mathbf{D}^i is a theorem of \mathcal{Q}_0 , since (as we shall see in Corollary 7302 below) one cannot effectively test whether an arbitrary wff is a theorem of \mathcal{Q}_0 , and it is essential that we be able to effectively test whether an alleged proof from hypotheses is in fact a proof from hypotheses. Therefore one must

supply a proof for any theorem one uses in a proof from hypotheses; this is the purpose of the sequence S_1 .

We now define $\mathcal{H} \vdash A$ to mean that there is a proof of A from the hypotheses \mathcal{H} . Note that $\emptyset \vdash A$ iff A is a theorem of Q_0 . Hence we write $\vdash A$ to mean that A is a theorem of Q_0 .

EXERCISES

X5100. Prove that $S_{A_\alpha}^{y_\alpha} S_{y_\alpha}^{x_\alpha} B_\beta = S_{A_\alpha}^{x_\alpha} B_\beta$ provided that y_α does not occur in B_β .

X5101. Prove that $S_{y_\beta}^{z_\beta} S_{z_\beta}^{x_\beta} A_\alpha = S_{y_\beta}^{x_\beta} A_\alpha$, provided that z_β does not occur free in A_α and z_β is free for x_β in A_α . Note explicitly where the conditions on the variables are used in your proof.

§52. Elementary Logic in Q_0

We shall now prove certain theorem schemata and derived rules of inference of Q_0 . Among other things, we shall show that appropriate analogues of theorems and rules of first-order logic can be derived in Q_0 .

5200. $\vdash A_\alpha = A_\alpha$

Proof:

$$\begin{array}{ll} (.1) \vdash [\lambda x_\alpha x_\alpha] A_\alpha = A_\alpha & \text{Axiom 4}_2 \\ (.2) \vdash A_\alpha = A_\alpha & \text{R: .1, .1} \end{array}$$

■

Note the special case: $\vdash T_o$

5201 Equality Rules (= Rules).

If $\mathcal{H} \vdash A$ and $\mathcal{H} \vdash A \equiv B$ then $\mathcal{H} \vdash B$.

If $\mathcal{H} \vdash A_\alpha = B_\alpha$ then $\mathcal{H} \vdash B_\alpha = A_\alpha$.

If $\mathcal{H} \vdash A_\alpha = B_\alpha$ and $\mathcal{H} \vdash B_\alpha = C_\alpha$ then $\mathcal{H} \vdash A_\alpha = C_\alpha$.

If $\mathcal{H} \vdash A_{\alpha\beta} = B_{\alpha\beta}$ and $\mathcal{H} \vdash C_\beta = D_\beta$ then $\mathcal{H} \vdash A_{\alpha\beta}C_\beta = B_{\alpha\beta}D_\beta$.

If $\mathcal{H} \vdash A_{\alpha\beta} = B_{\alpha\beta}$ then $\mathcal{H} \vdash A_{\alpha\beta}C_\beta = B_{\alpha\beta}C_\beta$.

If $\mathcal{H} \vdash C_\beta = D_\beta$ then $\mathcal{H} \vdash A_{\alpha\beta}C_\beta = A_{\alpha\beta}D_\beta$.

Proof: By 5200 and Rule R'. ■

5202 Rule R.R. If $\vdash A_\alpha = B_\alpha$ or $\vdash B_\alpha = A_\alpha$, and if $\mathcal{H} \vdash C$, then $\mathcal{H} \vdash D$, where D is obtained from C by replacing one occurrence of A_α in C by an occurrence of B_α , and the occurrence of A_α is not an occurrence of a variable immediately preceded by λ . (Note that since hypotheses were not needed to derive $A_\alpha = B_\alpha$, the restriction on Rule R' involving variables free in \mathcal{H} does not apply.)

Proof:

- | | |
|-----------------------------------|------------------------------|
| (.1) $\vdash A_\alpha = B_\alpha$ | given, and 5201 if necessary |
| (.2) $\vdash C = C$ | 5200 |
| (.3) $\vdash C = D$ | R |
| (.4) $\mathcal{H} \vdash C$ | given |
| (.5) $\mathcal{H} \vdash C = D$ | theorem |
| (.6) $\mathcal{H} \vdash D$ | R' |

■

5203. $\vdash [\lambda x_\alpha B_\beta] A_\alpha = S_{A_\alpha}^{x_\alpha} B_\beta$ provided no variable in A_α is bound in B_β .

Proof: By induction on the number of occurrences of [in B_β .

Case: B_β is a primitive constant or variable. Use 4₁ or 4₂.

Case: B_β has the form $[D_{\beta\gamma} C_\gamma]$.

- | | |
|--|-------------------------|
| $\vdash [\lambda x_\alpha D_{\beta\gamma}] A_\alpha = S_{A_\alpha}^{x_\alpha} D_{\beta\gamma}$ | by inductive hypothesis |
| $\vdash [\lambda x_\alpha C_\gamma] A_\alpha = S_{A_\alpha}^{x_\alpha} C_\gamma$ | by inductive hypothesis |
| $\vdash [\lambda x_\alpha \bullet D_{\beta\gamma} C_\gamma] A_\alpha = [[\lambda x_\alpha D_{\beta\gamma}] A_\alpha] [[\lambda x_\alpha C_\gamma] A_\alpha]$ | 4 ₃ |

Hence the desired theorem follows by Rule R.

Case: B_β has the form $[\lambda y_\gamma D_\delta]$. If y_γ is x_α , the result follows by 4₅. If y_γ is not x_α , then since y_γ is bound in B_β , it does not occur in A_α , so

- | | |
|--|--------------------------|
| $\vdash [\lambda x_\alpha \bullet \lambda y_\gamma D_\delta] A_\alpha = [\lambda y_\gamma \bullet [\lambda x_\alpha D_\delta] A_\alpha]$ | 4 ₄ |
| $\vdash [\lambda x_\alpha D_\delta] A_\alpha = S_{A_\alpha}^{x_\alpha} D_\delta$ | by inductive hypothesis. |

Hence the desired theorem follows by Rule R. ■

REMARK. Theorem Schema 5203 is a preliminary result, which will occur in more general form in 5207. We next prove a restricted substitution rule, which will occur in more general form in 5209, and in full generality in 5221.

5204. If $\vdash \mathbf{B}_\beta = \mathbf{C}_\beta$, then $\vdash S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} [\mathbf{B}_\beta = \mathbf{C}_\beta]$, provided no variable in \mathbf{A}_α is bound in \mathbf{B}_β or \mathbf{C}_β .

Proof:

- (.1) $\vdash [\lambda \mathbf{x}_\alpha \mathbf{B}_\beta] \mathbf{A}_\alpha = [\lambda \mathbf{x}_\alpha \mathbf{B}_\beta] \mathbf{A}_\alpha$ 5200
- (.2) $\vdash [\lambda \mathbf{x}_\alpha \mathbf{B}_\beta] \mathbf{A}_\alpha = [\lambda \mathbf{x}_\alpha \mathbf{C}_\beta] \mathbf{A}_\alpha$ R: .1
- (.3) $\vdash [\lambda \mathbf{x}_\alpha \mathbf{B}_\beta] \mathbf{A}_\alpha = S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} \mathbf{B}_\beta$ 5203
- (.4) $\vdash [\lambda \mathbf{x}_\alpha \mathbf{C}_\beta] \mathbf{A}_\alpha = S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} \mathbf{C}_\beta$ 5203
- (.5) $\vdash S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} [\mathbf{B}_\beta = \mathbf{C}_\beta]$ R: .2, .3, 4. ■

5205. $\vdash f_{\alpha\beta} = [\lambda \mathbf{y}_\beta \bullet f_{\alpha\beta} \mathbf{y}_\beta]$

Proof: First we handle the case where \mathbf{y}_β is distinct from x_β (the particular variable which is bound in Axiom 3 $^\alpha$).

- (.1) $\vdash f_{\alpha\beta} = [\lambda \mathbf{y}_\beta \bullet f_{\alpha\beta} \mathbf{y}_\beta] = \forall x_\beta \bullet f_{\alpha\beta} x_\beta = [\lambda \mathbf{y}_\beta \bullet f_{\alpha\beta} \mathbf{y}_\beta] x_\beta$ 5204: 3 $^\alpha$
- (.2) $\vdash [\lambda \mathbf{y}_\beta \bullet f_{\alpha\beta} \mathbf{y}_\beta] x_\beta = f_{\alpha\beta} x_\beta$ 5203
- (.3) $\vdash f_{\alpha\beta} = [\lambda \mathbf{y}_\beta \bullet f_{\alpha\beta} \mathbf{y}_\beta] = \forall x_\beta \bullet f_{\alpha\beta} x_\beta = f_{\alpha\beta} x_\beta$ R: .1, .2
- (.4) $\vdash f_{\alpha\beta} = f_{\alpha\beta} = \forall x_\beta \bullet f_{\alpha\beta} x_\beta = f_{\alpha\beta} x_\beta$ 5204: 3 $^\alpha$
- (.5) $\vdash f_{\alpha\beta} = [\lambda \mathbf{y}_\beta \bullet f_{\alpha\beta} \mathbf{y}_\beta]$ where \mathbf{y}_β is distinct from x_β .
= Rules: .3, .4, 5200

Next we handle the case where \mathbf{y}_β is x_β .

- (.6) $\vdash f_{\alpha\beta} = [\lambda y_\beta \bullet f_{\alpha\beta} y_\beta]$ instance of .5
- (.7) $\vdash [\lambda x_\beta \bullet f_{\alpha\beta} x_\beta] = [\lambda y_\beta \bullet [\lambda x_\beta \bullet f_{\alpha\beta} x_\beta] y_\beta]$ 5204: .6
- (.8) $\vdash [\lambda x_\beta \bullet f_{\alpha\beta} x_\beta] = [\lambda y_\beta \bullet f_{\alpha\beta} y_\beta]$ R: .7, 5203
- (.9) $\vdash f_{\alpha\beta} = [\lambda x_\beta \bullet f_{\alpha\beta} x_\beta]$ = Rules: .6, .8 ■

5206 $\vdash [\lambda \mathbf{x}_\beta \mathbf{A}_\alpha] = [\lambda \mathbf{z}_\beta S_{\mathbf{z}_\beta}^{\mathbf{x}_\beta} \mathbf{A}_\alpha]$, provided that \mathbf{z}_β does not occur free in \mathbf{A}_α and \mathbf{z}_β is free for \mathbf{x}_β in \mathbf{A}_α .

REMARK. A natural way to try to prove this is to start with

$$[\lambda \mathbf{x}_\beta \mathbf{A}_\alpha] = [\lambda \mathbf{z}_\beta \bullet [\lambda \mathbf{x}_\beta \mathbf{A}_\alpha] \mathbf{z}_\beta] \quad 5204: 5205$$

and then use 5203. However, this application of 5204 will not work if \mathbf{z}_β occurs (as a bound variable) in \mathbf{A}_α , so we use the method below.

Proof: Let y_β be distinct from x_β , z_β , and all variables in A_α .

- | | |
|---|-------------|
| (.1) $\vdash [\lambda x_\beta A_\alpha] = [\lambda y_\beta \bullet [\lambda x_\beta A_\alpha] y_\beta]$ | 5204: 5205 |
| (.2) $\vdash [\lambda x_\beta A_\alpha] = [\lambda y_\beta \bullet S_{y_\beta}^{x_\beta} A_\alpha]$ | R: .1, 5203 |
| (.3) $\vdash [\lambda z_\beta S_{z_\beta}^{x_\beta} A_\alpha] = [\lambda y_\beta \bullet [\lambda z_\beta S_{z_\beta}^{x_\beta} A_\alpha] y_\beta]$ | 5204: 5205 |
| (.4) $\vdash [\lambda z_\beta S_{z_\beta}^{x_\beta} A_\alpha] = [\lambda y_\beta \bullet S_{y_\beta}^{x_\beta} A_\alpha]$ | R: .3, 5203 |

(.4) is obtained as specified because $S_y^z S_z^x A_\alpha = S_y^x A_\alpha$, as shown in Exercise X5101. The theorem then follows from (.2) and (.4) by = Rules. ■

5207 $\vdash [\lambda x_\alpha B_\beta] A_\alpha = S_{A_\alpha}^{x_\alpha} B_\beta$, provided that A_α is free for x_α in B_β .

Proof: This theorem follows from 5203 and 5206. For a careful proof, we proceed as in the proof of 5203 by induction on the number of left brackets in B_β . The cases where B_β is a primitive constant or variable, has the form $[D_\beta C_\gamma]$, or has the form $[\lambda y_\gamma D_\delta]$ where y_γ is x_α , follow just as in the proof of 5203.

Hence we need deal only with the case where B_β has the form $[\lambda y_\gamma D_\delta]$, and y_γ is not x_α . Thus we must prove $\vdash [\lambda x_\alpha \bullet \lambda y_\gamma D_\delta] A_\alpha = [\lambda y_\gamma S_{A_\alpha}^{x_\alpha} D_\delta]$. If y_γ does not occur in A_α , we can use Axiom 4₄ and the inductive hypothesis. Of course, y_γ may actually occur in A_α . However, bound occurrences of y_γ can be changed to occurrences of some new variable z_γ , and this is done in both cases below.

Let z_γ be a variable distinct from x_α , y_γ , and all variables in A_α or D_δ .

Subcase. x_α is not free in D_δ . Let D'_δ be $S_{z_\gamma}^y D_\delta$. (We shall replace $[\lambda y_\gamma D_\delta]$ by $[\lambda z_\gamma D'_\delta]$ so that the restriction on Axiom 4₄ will be satisfied.)

- | | |
|---|---|
| (.1) $\vdash [\lambda x_\alpha [\lambda z_\gamma D'_\delta]] A_\alpha = [\lambda z_\gamma \bullet [\lambda x_\alpha D'_\delta] A_\alpha]$ | 4 ₄ |
| (.2) $\vdash [\lambda x_\alpha D'_\delta] A_\alpha = D'_\delta$ | by inductive hypothesis; the right side
is D'_δ since x_α is not free in D'_δ . |
| (.3) $\vdash [\lambda x_\alpha [\lambda z_\gamma D'_\delta]] A_\alpha = [\lambda z_\gamma D'_\delta]$ | R: .1, .2 |
| (.4) $\vdash [\lambda y_\gamma D_\delta] = [\lambda z_\gamma D'_\delta]$ | 5206 |
| (.5) $\vdash [\lambda x_\alpha [\lambda y_\gamma D_\delta]] A_\alpha = [\lambda y_\gamma D_\delta]$ | RR: .3, .4 |

This is the desired theorem for this subcase.

Subcase. x_α is free in D_δ . (Here we will change y_γ in A_α .) Since A_α is free for x_α in B_β , y_γ does not occur free in A_α . Let A'_α be the result of replacing all (bound) occurrences of y_γ in A_α by occurrences of z_γ .

- (.6) $\vdash A_\alpha = A'_\alpha \quad \text{R: 5200, 5206}$
 (Several different instances of 5206 may be needed.)
- (.7) $\vdash [\lambda x_\alpha \bullet \lambda y_\gamma D_\delta] A'_\alpha = [\lambda y_\gamma \bullet [\lambda x_\alpha D_\delta] A'_\alpha] \quad 4_4$
- (.8) $\vdash [\lambda x_\alpha \bullet \lambda y_\gamma D_\delta] A_\alpha = [\lambda y_\gamma \bullet [\lambda x_\alpha D_\delta] A_\alpha] \quad \text{RR: } .6, .7$
- (.9) $\vdash [\lambda x_\alpha D_\delta] A_\alpha = S_{A_\alpha}^{x_\alpha} D_\delta \quad \text{by inductive hypothesis}$
- (.10) $\vdash [\lambda x_\alpha \bullet \lambda y_\gamma D_\delta] A_\alpha = [\lambda y_\gamma S_{A_\alpha}^{x_\alpha} D_\delta] \quad \text{R: } .8, .9$

This is the desired theorem. ■

We now introduce some traditional terminology by defining the derived rules of inference listed below, which are obtained directly from 5206, 5207, and 5205 by using Rule RR and 5204.

- **α -Conversion (Alphabetic Change of Bound Variables) (α).** Replace a wf part $[\lambda x_\beta A_\alpha]$ of a wff by $[\lambda z_\beta S_{z_\beta}^{x_\beta} A_\alpha]$, provided that z_β does not occur free in A_α and z_β is free for x_β in A_α .
- **β -Contraction or β -Reduction.** Replace a wf part $[\lambda x_\alpha B_\beta] A_\alpha$ of a wff by $S_{A_\alpha}^{x_\alpha} B_\beta$, provided A_α is free for x_α in B_β .
- **β -Expansion.** Apply the inverse of β -Contraction.
- **β -Conversion (β).** Apply β -Contraction or β -Expansion.
- **η -Contraction or η -Reduction.** Replace a wf part $[\lambda y_\beta \bullet B_{\alpha\beta} y_\beta]$ of a wff by $B_{\alpha\beta}$, provided y_β does not occur free in $B_{\alpha\beta}$.
- **η -Expansion.** Apply the inverse of η -Contraction.
- **η -Conversion (η).** Apply η -Contraction or η -Expansion.
- **λ -Conversion (λ).** Apply any combination of α -Conversions, β -Conversions, and η -Conversions.

5208 $\vdash [\lambda x_{\alpha_1}^1 \dots \lambda x_{\alpha_n}^n B_\beta] x_{\alpha_1}^1 \dots x_{\alpha_n}^n = B_\beta$ where $n \geq 1$.

Proof: By induction on n . For $n = 1$, the theorem is an instance of 5207.

Suppose all instances of 5208 can be proved for $n = m$; we must prove it for $n = m + 1$. (Note that we do not require the variables $x_{\alpha_i}^i$ to be distinct.)

$$(1) \vdash [\lambda x^1 \lambda x^2 \dots \lambda x^{m+1} B] x^1 x^2 \dots x^{m+1} = [\lambda x^2 \dots \lambda x^{m+1} B] x^2 \dots x^{m+1}$$

by 5200, 5207

(.2) $\vdash [\lambda x^2 \dots \lambda x^{m-1} B]x^2 \dots x^{m-1} = B$ by inductive hypothesis
 The theorem follows by Rule R. ■

5209. If $\vdash B_\beta = C_\beta$, then $\vdash S_{A_\alpha}^{x_\alpha}[B_\beta = C_\beta]$, provided A_α is free for x_α in $[B_\beta = C_\beta]$.

The proof follows as for 5204, with 5207 used in place of 5203.

5210 $\vdash T = [B_\beta = B_\beta]$

Proof:

- (.1) $\vdash [\lambda y_\beta y_\beta] = [\lambda y_\beta y_\beta] = \forall x_\beta \bullet [\lambda y_\beta y_\beta]x_\beta = [\lambda y_\beta y_\beta]x_\beta$ 5209: 3 $^{\beta\beta}$
- (.2) $\vdash [\lambda x_\beta T] = [\lambda x_\beta \bullet x_\beta = x_\beta]$ R: 5200, .1; λ , def of \forall
- (.3) $\vdash [\lambda x_\beta T]B_\beta = [\lambda x_\beta \bullet x_\beta = x_\beta]B_\beta$ =: .2
- (.4) $T = \bullet B_\beta = B_\beta$ λ : .3 ■

5211. $\vdash [T \wedge T] = T$

Proof:

- (.1) $\vdash [\lambda y_o T]T \wedge [\lambda y_o T]F = \forall x_o [\lambda y_o T]x_o$ 5209: Axiom 1
- (.2) $\vdash T \wedge T = \forall x_o T$ λ : .1
- (.3) $\vdash T = \forall x_o T$ 5210, def of \forall
- (.4) $\vdash T \wedge T = T$ =: .2, .3 ■

REMARK. An alternative proof can be given which does not use Axiom 1, but does use the definition of \wedge .

5212. $\vdash T \wedge T$

Proof: by = Rules (5201) from 5211, 5200, and the definition of T . Alternatively, by 5200, the definition of \wedge , and λ .

5213. If $\vdash A_\alpha = B_\alpha$ and $\vdash C_\beta = D_\beta$ then $\vdash [A_\alpha = B_\alpha] \wedge [C_\beta = D_\beta]$.

Proof:

- (.1) $\vdash A_\alpha = B_\alpha$ given
- (.2) $\vdash T = [A_\alpha = B_\alpha]$ R: 5210, .1
- (.3) $\vdash C_\beta = D_\beta$ given

- (.4) $\vdash T = [\mathbf{C}_\beta = \mathbf{D}_\beta]$
(.5) $\vdash [\mathbf{A}_\alpha = \mathbf{B}_\alpha] \wedge [\mathbf{C}_\beta = \mathbf{D}_\beta]$

R: 5210, .3
R: 5212, .2, .4 ■

5214. $\vdash T \wedge F = F$

Proof:

- (.1) $\vdash [\lambda x_o x_o]T \wedge [\lambda x_o x_o]F = \forall x_o [\lambda x_o x_o]x_o$
(.2) $\vdash T \wedge F = \forall x_o x_o$

5209: Axiom 1
λ: .1

By the definitions of \forall and F , we see that this is the desired theorem. ■

5215 Universal Instantiation ($\forall I$). If $\mathcal{H} \vdash \forall x_\alpha B$, then $\mathcal{H} \vdash S_{A_\alpha}^{x_\alpha} B$, provided that A_α is free for x_α in B .

Proof:

- (.1) $\mathcal{H} \vdash [\lambda x_\alpha T] = [\lambda x_\alpha B]$
(.2) $\mathcal{H} \vdash [\lambda x_\alpha T]A_\alpha = [\lambda x_\alpha B]A_\alpha$
(.3) $\mathcal{H} \vdash T = S_{A_\alpha}^{x_\alpha} B$
(.4) $\mathcal{H} \vdash S_{A_\alpha}^{x_\alpha} B$

given, def of \forall
=: .1
λ: .2
R: 5200, .3 ■

5216. $\vdash [T \wedge A] = A$

Proof:

- (.1) $\vdash [\lambda x_o \bullet T \wedge x_o = x_o]T \wedge [\lambda x_o \bullet T \wedge x_o = x_o]F$
 $= \forall x_o \bullet [\lambda x_o \bullet T \wedge x_o = x_o]x_o$
(.2) $\vdash [T \wedge T = T] \wedge [T \wedge F = F] = \forall x_o \bullet T \wedge x_o = x_o$
(.3) $\vdash [T \wedge T = T] \wedge [T \wedge F = F]$
(.4) $\vdash \forall x_o \bullet T \wedge x_o = x_o$
(.5) $\vdash T \wedge A = A$

5209: Axiom 1
λ: .1
5213: 5211, 5214
R: .2, .3
∀I: .4 ■

5217. $\vdash [T = F] = F$

Proof:

- (.1) $\vdash [\lambda x_o \bullet T = x_o]T \wedge [\lambda x_o \bullet T = x_o]F = \forall x_o \bullet [\lambda x_o \bullet T = x_o]x_o$
(.2) $\vdash [T = T] \wedge [T = F] = \forall x_o \bullet T = x_o$

5209: Axiom 1
λ: .1

- (.3) $\vdash T \wedge [T = F] = \forall x_o \bullet T = x_o$ RR: 5210, .2
(.4) $\vdash T = F = \forall x_o \bullet T = x_o$ R: 5216, .3
(.5) $\vdash [\lambda x_o T] = [\lambda x_o x_o] = \forall x_o \bullet [\lambda x_o T] x_o = [\lambda x_o x_o] x_o$ 5209: Axiom 3^{oo}
(.6) $\vdash F = \forall x_o \bullet T = x_o$ λ: .5; def of F
(.7) $\vdash [T = F] = F$ =: .4, .6 ■

5218 $\vdash [T = \mathbf{A}] = \mathbf{A}$

Proof:

- (.1) $\vdash [[T = T] = T] \wedge [[T = F] = F] = \forall x_o \bullet [T = x_o] = x_o$ 5209: Axiom 1; λ
(.2) $\vdash [T = T] = T$ =: 5210
(.3) $\vdash [T = T = T] \wedge [T = F = F]$ 5213: .2, 5217
(.4) $\vdash \forall x_o \bullet T = x_o = x_o$ R: .3, .1
(.5) $\vdash T = \mathbf{A} = \mathbf{A}$ ∀I: .3 ■

5219 Rule T. $\mathcal{H} \vdash \mathbf{A}$ if and only if $\mathcal{H} \vdash T = \mathbf{A}$; $\mathcal{H} \vdash \mathbf{A}$ if and only if $\mathcal{H} \vdash \mathbf{A} = T$.

Proof: By RR(5202), 5218, and =(5201). ■

5220 Rule of Universal Generalization (Gen). If $\mathcal{H} \vdash \mathbf{A}$ then $\mathcal{H} \vdash \forall x_\alpha \mathbf{A}$, provided that x_α is not free in any wff in \mathcal{H} .

Proof:

- (.1) $\mathcal{H} \vdash \mathbf{A}$ given
(.2) $\mathcal{H} \vdash T = \mathbf{A}$ Rule T: .1
(.3) $\mathcal{H} \vdash [\lambda x_\alpha T] = [\lambda x_\alpha T]$ 5206
(.4) $\mathcal{H} \vdash \forall x_\alpha \mathbf{A}$ R': .2, .3; def of \forall

Note that the restrictions on Rule R' are satisfied. ■

5221n Rule of Substitution (Sub). If $\mathcal{H} \vdash \mathbf{B}$, then $\mathcal{H} \vdash S_{\mathbf{A}_{\alpha_1}^1 \dots \mathbf{A}_{\alpha_n}^n}^{x_{\alpha_1}^1 \dots x_{\alpha_n}^n} \mathbf{B}$, provided that the variables x^1, \dots, x^n are distinct from one another and from all free variables of wffs in \mathcal{H} , and that $\mathbf{A}_{\alpha_i}^i$ is free for $x_{\alpha_i}^i$ in \mathbf{B} for all i ($1 \leq i \leq n$).

Proof:

First we consider the case $n = 1$.

- (.1) $\mathcal{H} \vdash \mathbf{B}$ given
- (.2) $\mathcal{H} \vdash \forall x_\alpha \mathbf{B}$ Gen: .1
- (.3) $\mathcal{H} \vdash S_{A_\alpha}^{x_\alpha} \mathbf{B}$. $\forall I(5215)$: .2

Next consider the case of an arbitrary number n . Let $y_{\alpha_1}^1, \dots, y_{\alpha_n}^n$ be distinct from one another, from x^1, \dots, x^n , and from all variables in \mathbf{B} , $A_{\alpha_1}^1, \dots, A_{\alpha_n}^n$, and wffs in \mathcal{H} .

- (.4) $\mathcal{H} \vdash S_{y_1 \dots y_n}^{x_1 \dots x_n} \mathbf{B}$ 5221₁ (applied n times)
- (.5) $\mathcal{H} \vdash S_{A_1 \dots A_n}^{y_1 \dots y^n} [S_{y_1 \dots y^n}^{x_1 \dots x^n} \mathbf{B}]$ 5221₁ (applied n times): .4

This is the desired wff. ■

Since we have now established that we can substitute freely for variables of Q_0 , we can henceforth indicate that all wffs of Q_0 having a certain form are theorems in either of two ways: by stating a theorem schema (such as $T = \mathbf{A} = \mathbf{A}$), or by stating one instance of it (such as $T = x_o = x_o$) from which all others can be obtained by substitution. In the former case we use syntactical variables from our meta-language, and in the latter we use variables from the object language Q_0 . We shall generally prefer to use the variables of Q_0 .

When no confusion seems likely, we shall often use the number of a theorem to designate any theorem which can be derived from it by the α and Substitution Rules, and omit explicit reference to these rules in justifying the steps of a proof.

5222 Rule of Cases. If $\mathcal{H} \vdash S_T^{x_o} \mathbf{A}$ and $\mathcal{H} \vdash S_F^{x_o} \mathbf{A}$, then $\mathcal{H} \vdash \mathbf{A}$.

Proof:

- (.1) $\mathcal{H} \vdash T = [\lambda x_o \mathbf{A}]T$ Rule T, λ : hypothesis
- (.2) $\mathcal{H} \vdash T = [\lambda x_o \mathbf{A}]F$ Rule T, λ : hypothesis
- (.3) $\mathcal{H} \vdash T \wedge T$ 5212
- (.4) $\mathcal{H} \vdash [\lambda x_o \mathbf{A}]T \wedge [\lambda x_o \mathbf{A}]F$ R': .1, .2, .3
- (.5) $\vdash [\lambda x_o \mathbf{A}]T \wedge [\lambda x_o \mathbf{A}]F = \forall x_o \mathbf{A}$ α , Sub, λ : Axiom 1
- (.6) $\mathcal{H} \vdash \forall x_o \mathbf{A}$ RR: .4, .5
- (.7) $\mathcal{H} \vdash \mathbf{A}$ $\forall I(5215)$: .6 ■

5223. $\vdash [T \supset y_o] = y_o$

Proof:

- (.1) $\vdash [T \supset y] = \blacksquare T = [T \wedge y]$ λ , def of \supset
 - (.2) $\vdash [T \supset y] = \blacksquare T \wedge y$ R: .1, 5218
 - (.3) $\vdash [T \supset y] = y$ R: .2, 5216
-

5224 Modus Ponens (MP). If $\mathcal{H} \vdash A$ and $\mathcal{H} \vdash A \supset B$, then $\mathcal{H} \vdash B$.

Proof:

- (.1) $\mathcal{H} \vdash A \supset B$ given
 - (.2) $\mathcal{H} \vdash A = T$ Rule T: hypothesis
 - (.3) $\mathcal{H} \vdash T \supset B$ $R': .1, .2$
 - (.4) $\mathcal{H} \vdash B$ RR: .3, 5223
-

5225. $\vdash \Pi_{o(o\alpha)} f_{o\alpha} \supset f_{o\alpha} x_\alpha$

Proof:

- (.1) $\vdash \Pi_{o(o\alpha)} f_{o\alpha} \supset \blacksquare [\lambda f_{o\alpha} \cdot f_{o\alpha} x_\alpha][\lambda x_\alpha T] = [\lambda f_{o\alpha} \cdot f_{o\alpha} x_\alpha] f_{o\alpha}$ Sub: Axiom $2^{o\alpha}$; def of $\Pi_{o(o\alpha)}$
 - (.2) $\vdash \Pi_{o(o\alpha)} f_{o\alpha} \supset \blacksquare T = f_{o\alpha} x_\alpha$ $\lambda: .1$
 - (.3) $\vdash \Pi_{o(o\alpha)} f_{o\alpha} \supset f_{o\alpha} x_\alpha$ R: 5218, .2
-
- (Note that this is our first use of Axiom 2.)

5226. $\vdash \forall x_\alpha B \supset S_{A_\alpha}^{x_\alpha} B$ provided that A_α is free for x_α in B .

Proof:

- (.1) $\vdash \Pi_{o(o\alpha)} [\lambda x_\alpha B] \supset [\lambda x_\alpha B] A_\alpha$ Sub: 5225
 - (.2) $\vdash \forall x_\alpha B \supset S_{A_\alpha}^{x_\alpha} B$ $\lambda: .1$
-

5227. $\vdash F \supset x_o$

This is a special case of 5226, since F_o is $\forall x_o x_o$.

5228. $\vdash [T \supset T] = T; \vdash [T \supset F] = F; \vdash [F \supset T] = T; \vdash [F \supset F] = T$

Proof: by 5223, 5227, and Rule T. ■

5229. $\vdash [T \wedge T] = T; \vdash [T \wedge F] = F; \vdash [F \wedge T] = F; \vdash [F \wedge F] = F$

Proof: $\vdash F \wedge A = F$ by 5227, the definition of \supset , and $=$ Rules. Also use 5216. ■

5230. $\vdash [T = T] = T; \vdash [T = F] = F; \vdash [F = T] = F; \vdash [F = F] = T$

Proof: From 5210 and 5218 one obtains all of these except $[F = T] = F$; to prove this we proceed as follows:

- (.1) $\vdash [F = T] \supset \bullet [\lambda x_o \bullet x_o = F]F = [\lambda x_o \bullet x_o = F]T$ Sub: Axiom 2⁰
- (.2) $\vdash [F = T] \supset \bullet [F = F] = [T = F]$ λ: .1
- (.3) $\vdash [F = T] \supset F$ RR: .2, 5210, 5218
- (.4) $\vdash [F = T] = \bullet [F = T] \wedge F$ λ: .3, def of \supset
- (.5) $\vdash x_o \wedge F = F$ Rule of Cases (5222): 5229
- (.6) $\vdash [F = T] \wedge F = F$ Sub: .5
- (.7) $\vdash [F = T] = F$ =: .4, .6

5231. $\vdash \sim T = F; \vdash \sim F = T$

Proof: By 5230 and the definition of \sim . ■

5232. $\vdash T \vee T = T; \vdash T \vee F = T; \vdash F \vee T = T; \vdash F \vee F = F$

Proof: By 5229, 5231, λ, and the definition of \vee_{ooo} . ■

DEFINITION. The class of *propositional wffs* is the smallest class of wffs_o which contains T_o , F_o , all variables of type o, and which, if it contains **A** and **B**, also contains $[\sim A]$, $[A \wedge B]$, $[A \vee B]$, $[A \supset B]$, and $[A = B]$. Given any assignment φ of truth values to variables of type o and a propositional wff **A**, we define $\mathcal{V}_\varphi A$ (*the truth value of A with respect to φ*) as follows by induction on the number of occurrences of [in **A**:

- (1) $\mathcal{V}_\varphi T_o = T$ (truth)
- (2) $\mathcal{V}_\varphi F_o = F$ (falsehood)
- (3) $\mathcal{V}_\varphi p_o = \varphi p_o$
- (4) $\mathcal{V}_\varphi \sim A = T$ if $\mathcal{V}_\varphi A = F$
 $= F$ if $\mathcal{V}_\varphi A = T$
- (5) $\mathcal{V}_\varphi [A \wedge B] = T$ if $\mathcal{V}_\varphi A = \mathcal{V}_\varphi B = T$
 $= F$ otherwise

$$(6) \quad \mathcal{V}_\varphi[\mathbf{A} \vee \mathbf{B}] = \begin{cases} \mathbf{T} & \text{if } \mathcal{V}_\varphi \mathbf{A} = \mathbf{T} \text{ or } \mathcal{V}_\varphi \mathbf{B} = \mathbf{T} \\ \mathbf{F} & \text{otherwise} \end{cases}$$

$$(7) \quad \mathcal{V}_\varphi[\mathbf{A} \supset \mathbf{B}] = \begin{cases} \mathbf{T} & \text{if } \mathcal{V}_\varphi \mathbf{A} = \mathbf{F} \text{ or } \mathcal{V}_\varphi \mathbf{B} = \mathbf{T} \\ \mathbf{F} & \text{otherwise} \end{cases}$$

$$(8) \quad \mathcal{V}_\varphi[\mathbf{A} = \mathbf{B}] = \begin{cases} \mathbf{T} & \text{if } \mathcal{V}_\varphi \mathbf{A} = \mathcal{V}_\varphi \mathbf{B} \\ \mathbf{F} & \text{otherwise} \end{cases}$$

REMARK. If \mathbf{A} is a propositional wff without free variables, then $\mathcal{V}_\varphi \mathbf{A} = \mathcal{V}_\psi \mathbf{A}$ for all assignments φ and ψ , so we let $\mathcal{V}\mathbf{A} = \mathcal{V}_\varphi \mathbf{A}$ for any assignment φ .

DEFINITION. A wff \mathbf{A} is a *tautology* iff \mathbf{A} is a propositional wff and $\mathcal{V}_\varphi \mathbf{A} = \mathbf{T}$ for all truth-value assignments φ . A wff \mathbf{B} is a *substitution instance of a tautology*, or is *tautologous*, iff there exists a tautology \mathbf{A} , wffs $\mathbf{C}^1, \dots, \mathbf{C}^n$, and distinct variables $\mathbf{p}_o^1, \dots, \mathbf{p}_o^n$ such that $\mathbf{B} = S_{\mathbf{C}^1 \dots \mathbf{C}^n}^{\mathbf{p}_o^1 \dots \mathbf{p}_o^n} \mathbf{A}$.

5233. If \mathbf{A} is a tautology, then $\vdash \mathbf{A}$.

Proof: By induction on the number n of free variables in \mathbf{A} .

Case: $n = 0$

We prove by induction on the number of occurrences of [in \mathbf{B} that if \mathbf{B} is a propositional wff without free variables, then $\vdash \mathbf{B} = \mathbf{T}$ if $\mathcal{V}\mathbf{B} = \mathbf{T}$, and $\vdash \mathbf{B} = \mathbf{F}$ if $\mathcal{V}\mathbf{B} = \mathbf{F}$.

Subcase. \mathbf{B} is \mathbf{T} or \mathbf{F} . Use 5200.

Subcase. \mathbf{B} has the form $[\sim \mathbf{C}]$.

If $\mathcal{V}\mathbf{C} = \mathbf{F}$ then $\vdash \mathbf{C} = \mathbf{F}$ by inductive hypothesis, so $\vdash \sim \mathbf{C} = \mathbf{T}$ by 5231 and = Rules. Also $\mathcal{V}\mathbf{B} = \mathbf{T}$.

If $\mathcal{V}\mathbf{C} = \mathbf{T}$ then $\vdash \mathbf{B} = \mathbf{F}$ and $\mathcal{V}\mathbf{B} = \mathbf{F}$ by a similar argument.

The cases where \mathbf{B} is a conjunction, disjunction, implication, or equivalence are handled similarly using 5229, 5232, 5228, and 5230, respectively.

Induction step. If \mathbf{A} is a tautology with at least one free variable \mathbf{p}_o , then $S_T^{\mathbf{p}_o} \mathbf{A}$ and $S_F^{\mathbf{p}_o} \mathbf{A}$ are tautologies with fewer free variables than \mathbf{A} , so by the inductive hypothesis (and Rule T when $n = 1$) $\vdash S_T^{\mathbf{p}_o} \mathbf{A}$ and $\vdash S_F^{\mathbf{p}_o} \mathbf{A}$. Therefore $\vdash \mathbf{A}$ by the Rule of Cases (5222). ■

5234 Rule P. If $\mathcal{H} \vdash A^1, \dots, \mathcal{H} \vdash A^n$, and if $[A^1 \wedge \dots \wedge A^n] \supset B$ is tautologous, then $\mathcal{H} \vdash B$. Also, if B is tautologous, then $\mathcal{H} \vdash B$.

Proof: Clearly $A^1 \supset \dots \supset A^n \supset B$ is also tautologous, and therefore a theorem by Rule Sub (5221) and 5233. Hence $\mathcal{H} \vdash A^1 \supset \dots \supset A^n \supset B$. Then $\mathcal{H} \vdash B$ by n applications of Modus Ponens (5224). ■

5235. $\vdash \forall x_\alpha [A \vee B] \supset \neg A \vee \forall x_\alpha B$ provided x_α is not free in A .

Proof: Let p_o be a variable which does not occur in the wff to be proved.

- | | |
|--|------------------------------|
| (.1) $\vdash \forall x_\alpha [T \vee B] \supset \neg T \vee \forall x_\alpha B$ | Rule P |
| (.2) $\vdash \forall x_\alpha B \supset \neg F \vee \forall x_\alpha B$ | Rule P |
| (.3) $\vdash B = \neg F \vee B$ | Rule P |
| (.4) $\vdash \forall x_\alpha [F \vee B] \supset \neg F \vee \forall x_\alpha B$ | R: .2, .3 |
| (.5) $\vdash \forall x_\alpha [p_o \vee B] \supset \neg p_o \vee \forall x_\alpha B$ | Rule of Cases (5222): .1, .4 |
| (.6) $\vdash \forall x_\alpha [A \vee B] \supset \neg A \vee \forall x_\alpha B$ | Sub: .5 |

■

REMARK. In view of 5220 (Gen), 5234 (Rule P), 5226, and 5235, it is now clear that all substitution instances of theorems of quantification theory (first-order logic) are theorems of Q_0 . We may sometimes make use of this fact by citing the following metatheorem to indicate informally how certain inferences in an abbreviated proof can be justified.

5236 “Rule Q”. If $\mathcal{H} \vdash A^1, \dots$, and $\mathcal{H} \vdash A^n$ (where $n \geq 0$), and if \bar{A}^i can be obtained from A^i by generalizing on zero or more free variables of A^i which do not occur free in \mathcal{H} , and if $[\bar{A}^1 \wedge \dots \wedge \bar{A}^n \supset B]$ is a substitution instance of a theorem of quantification theory, then $\mathcal{H} \vdash B$.

REMARK. Rule Q is not an effective rule of inference, since there is no effective procedure for determining whether an alleged application of Rule Q is in fact an application of the rule. Of course one can always cite the particular theorem of quantification theory which one is using in applying Rule Q, and supply a proof of that theorem. However, we will normally use Rule Q only when it will be rather obvious which well-known theorem of quantification theory is being used.

5237 $\supset \forall$ Rule. If $\mathcal{H} \vdash A \supset B$, then $\mathcal{H} \vdash A \supset \forall x B$, provided that x_α is not free in A or in \mathcal{H} .

Proof: by Rule Q, using Axiom 5 of \mathcal{F} . ■

5238. $\vdash [\lambda x_{\beta_1}^1 \dots \lambda x_{\beta_n}^n A_\alpha] = [\lambda x_{\beta_1}^1 \dots \lambda x_{\beta_n}^n B_\alpha] = \forall x_{\beta_1}^1 \dots \forall x_{\beta_n}^n [A_\alpha = B_\alpha]$

Proof: By induction on n .

$$(.) \vdash [\lambda x_\beta A_\alpha] = [\lambda x_\beta B_\alpha] = \forall x_\beta \bullet A_\alpha = B_\alpha \quad \alpha, \text{ Sub}, \lambda: \text{Axiom } 3^{\alpha\beta}$$

This takes care of the case $n = 1$. For the induction step we proceed as follows:

$$(.) \vdash [\lambda x^1 \dots \lambda x^{n-1} [\lambda x^n A_\alpha]] = [\lambda x^1 \dots \lambda x^{n-1} [\lambda x^n B_\alpha]] \\ = \forall x^1 \dots \forall x^{n-1} \bullet [\lambda x^n A_\alpha] = [\lambda x^n B_\alpha] \quad \text{by inductive hypothesis}$$

The desired result follows from .2 and an appropriate instance of .1 by Rule R. ■

5239. $\vdash \forall x_{\beta_1}^1 \dots \forall x_{\beta_n}^n [A_\alpha = B_\alpha] \supset \bullet C = D$, where $n \geq 0$ and

- (a) D is obtained from C and $[A_\alpha = B_\alpha]$ as in Rule R;
- (b) $x_{\beta_1}^1, \dots, x_{\beta_n}^n$ is a complete list of those variables x_β such that x_β occurs free in $[A_\alpha = B_\alpha]$, and the occurrence of A_α in C (which is replaced by B_α in D) is in a wf part of C of the form $[\lambda x_\beta E_\gamma]$.

Proof: Let $t_{\alpha\beta_n\dots\beta_1}$ be a variable which does not occur in the wff to be proved, and let G_o be the result of replacing the occurrence A_α in C by an occurrence of $[t_{\alpha\beta_n\dots\beta_1} x_{\beta_1}^1 \dots x_{\beta_n}^n]$.

$$(.) \vdash [[\lambda x^1 \dots \lambda x^n A_\alpha] = [\lambda x^1 \dots \lambda x^n B_\alpha]] \\ \supset \bullet [\lambda t_{\alpha\beta_n\dots\beta_1} G][\lambda x^1 \dots \lambda x^n A_\alpha] = [\lambda t_{\alpha\beta_n\dots\beta_1} G][\lambda x^1 \dots \lambda x^n B_\alpha] \\ \text{Sub: Axiom } 2^{(\alpha\beta_n\dots\beta_1)}$$

The desired theorem follows by 5238 and λ (5207, 5208). Note that the case where $n = 0$ presents no difficulties. ■

NOTATION. $\mathcal{H}, H \vdash P$ means $\mathcal{H} \cup \{H\} \vdash P$.

5240 Deduction Theorem. If $\mathcal{H}, H \vdash P$ then $\mathcal{H} \vdash H \supset P$.

Proof: Let S_1, S_2 be a proof of P from the hypotheses \mathcal{H}, H . We prove by induction on i that $\mathcal{H} \vdash H \supset R^i$ for each wff R^i in the sequence S_2 .

Case: R^i is a theorem or H .

- (.1) $\vdash \mathbf{H} \supset \mathbf{R}^i$
(.2) $\mathcal{H} \vdash \mathbf{H} \supset \mathbf{R}^i$

Rule P
theorem

Case: $\mathbf{R}^i \in \mathcal{H}$.

- (.3) $\mathcal{H} \vdash \mathbf{R}^i$
(.4) $\mathcal{H} \vdash \mathbf{H} \supset \mathbf{R}^i$

hypothesis
Rule P: .3

Case: \mathbf{R}^i is obtained from \mathbf{C} and $[\mathbf{A}_\alpha = \mathbf{B}_\alpha]$ by Rule \mathbf{R}' .

- (.5) $\mathcal{H} \vdash \mathbf{H} \supset \mathbf{C}$
(.6) $\mathcal{H} \vdash \mathbf{H} \supset \blacksquare \mathbf{A}_\alpha = \mathbf{B}_\alpha$

by inductive hypothesis
by inductive hypothesis

Let $\mathbf{x}_{\beta_1}^1, \dots, \mathbf{x}_{\beta_n}^n$ be a complete list of those variables \mathbf{x}_β such that \mathbf{x}_β occurs free in $[\mathbf{A}_\alpha = \mathbf{B}_\alpha]$ but the occurrence of \mathbf{A}_α in \mathbf{C} is in a wf part $[\lambda \mathbf{x}_\beta \mathbf{E}_\gamma]$ of \mathbf{C} . Note that the condition on Rule \mathbf{R}' assures that none of these variables is free in \mathcal{H} or \mathbf{H} .

- (.7) $\mathcal{H} \vdash \mathbf{H} \supset \forall \mathbf{x}_{\beta_1}^1 \dots \forall \mathbf{x}_{\beta_n}^n \blacksquare \mathbf{A}_\alpha = \mathbf{B}_\alpha$
(.8) $\mathcal{H} \vdash \mathbf{H} \supset \mathbf{R}^i$

5237: .6
Rule P: .5, .7, 5239

5241. If $\mathcal{H} \vdash \mathbf{A}$ and $\mathcal{H} \subseteq \mathcal{G}$ then $\mathcal{G} \vdash \mathbf{A}$.

Proof: This is clear when $\mathcal{H} = \emptyset$; if $\mathcal{H} \neq \emptyset$, let $\mathcal{H} = \{\mathbf{H}^1, \dots, \mathbf{H}^n\}$.

- (.1) $\vdash \mathbf{H}^1 \supset \blacksquare \dots \supset \blacksquare \mathbf{H}^n \supset \mathbf{A}$
(.2ⁱ) $\mathcal{G} \vdash \mathbf{H}^i$ for $1 \leq i \leq n$
(.3) $\mathcal{G} \vdash \mathbf{A}$

5240
hypothesis
Rule P

DEFINITION. If \mathcal{G} is an infinite set of wffs_o, we define $\mathcal{G} \vdash A$ to mean there is a finite subset \mathcal{H} of \mathcal{G} such that $\mathcal{H} \vdash A$. Note that 5240 and 5241 still apply when \mathcal{H} is infinite.

5242 Rule of Existential Generalization (\exists Gen). If $\mathcal{H} \vdash S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} \mathbf{B}$, and \mathbf{A}_α is free for \mathbf{x}_α in \mathbf{B} , then $\mathcal{H} \vdash \exists \mathbf{x}_\alpha \mathbf{B}$.

Proof: by Rule P, 5226, and the definition of \exists .

5243 Comprehension Theorem.

$\vdash \exists \mathbf{u}_{\beta \alpha_n \dots \alpha_1} \forall \mathbf{x}_{\alpha_1}^1 \dots \forall \mathbf{x}_{\alpha_n}^n \blacksquare \mathbf{u}_{\beta \alpha_n \dots \alpha_1} \mathbf{x}_{\alpha_1}^1 \dots \mathbf{x}_{\alpha_n}^n = \mathbf{B}_\beta$
when $n \geq 0$ and $\mathbf{u}_{\beta \alpha_n \dots \alpha_1}$ is not free in \mathbf{B}_β .

Proof:

$$(.) \vdash \forall x_{\alpha_1}^1 \dots \forall x_{\alpha_n}^n \bullet [\lambda x_{\alpha_1}^1 \dots \lambda x_{\alpha_n}^n B_\beta] x_{\alpha_1}^1 \dots x_{\alpha_n}^n = B_\beta$$

Gen: 5208. (Use 5200 when $n = 0$.)

Then use Existential Generalization. ■

5244 Existential Rule (\exists Rule). If $\mathcal{H}, B \vdash A$, and x_α is not free in \mathcal{H} or in A , then $\mathcal{H}, \exists x_\alpha B \vdash A$.

Proof:

- | | |
|--|---|
| (.1) $\mathcal{H} \vdash B \supset A$
(.2) $\mathcal{H} \vdash \exists x_\alpha B \supset A$
(.3) $\mathcal{H}, \exists x_\alpha B \vdash \exists x_\alpha B \supset A$
(.4) $\mathcal{H}, \exists x_\alpha B \vdash A$ | Deduction Theorem
Rule Q: .1
5241: .2
MP: .3, hypothesis |
|--|---|

5245 Rule C. If $\mathcal{H} \vdash \exists x_\alpha B$ and $\mathcal{H}, S_{y_\alpha}^{x_\alpha} B \vdash A$, where y_α is free for x_α in B and y_α is not free in $\mathcal{H}, \exists x_\alpha B$, or A , then $\mathcal{H} \vdash A$.

Proof:

- | | |
|--|--|
| (.1) $\mathcal{H}, S_{y_\alpha}^{x_\alpha} B \vdash A$
(.2) $\mathcal{H}, \exists y_\alpha [S_{y_\alpha}^{x_\alpha} B] \vdash A$
(.3) $\mathcal{H} \vdash \exists y_\alpha [S_{y_\alpha}^{x_\alpha} B] \supset A$
(.4) $\mathcal{H} \vdash \exists x_\alpha B \supset A$
(.5) $\mathcal{H} \vdash \exists x_\alpha B$
(.6) $\mathcal{H} \vdash A$ | given
\exists Rule: .1
Deduction Theorem: .2
α : .3
given
MP: .4, .5 |
|--|--|

DEFINITIONS.

$\subseteq_{o(o\alpha)(o\alpha)}$	stands for $[\lambda x_{o\alpha} \lambda y_{o\alpha} \forall z_\alpha \cdot xz \supset yz]$ (subset)
$\mathbf{A}_{o\alpha} \subseteq \mathbf{B}_{o\alpha}$	stands for $\subseteq_{o(o\alpha)(o\alpha)} \mathbf{A}_{o\alpha} \mathbf{B}_{o\alpha}$
$\mathcal{P}_{(o(o\alpha))(o\alpha)}$	stands for $[\lambda y_{o\alpha} \lambda x_{o\alpha} \cdot x_{o\alpha} \subseteq y_{o\alpha}]$ (power set)
$\cup_{o\alpha(o\alpha)(o\alpha)}$	stands for $[\lambda x_{o\alpha} \lambda y_{o\alpha} \lambda z_\alpha \cdot xz \vee yz]$ (union)
$\mathbf{A}_{o\alpha} \cup \mathbf{B}_{o\alpha}$	stands for $\cup_{o\alpha(o\alpha)(o\alpha)} \mathbf{A}_{o\alpha} \mathbf{B}_{o\alpha}$
$\bigcup_{o\alpha(o(o\alpha))}$	stands for $[\lambda w_{o(o\alpha)} \lambda z_\alpha \cdot \exists x_{o\alpha} \cdot wx \wedge xz]$ (union of a collection)
$\cap_{o\alpha(o\alpha)(o\alpha)}$	stands for $[\lambda x_{o\alpha} \lambda y_{o\alpha} \lambda z_\alpha \cdot xz \wedge yz]$ (intersection)
$\mathbf{A}_{o\alpha} \cap \mathbf{B}_{o\alpha}$	stands for $\cap_{o\alpha(o\alpha)(o\alpha)} \mathbf{A}_{o\alpha} \mathbf{B}_{o\alpha}$
$\bigcap_{o\alpha(o(o\alpha))}$	stands for $[\lambda w_{o(o\alpha)} \lambda z_\alpha \cdot \forall x_{o\alpha} \cdot wx \supset xz]$ (intersection of a collection)
$\#_{o\alpha(o\beta)(\alpha\beta)}$	stands for $[\lambda f_{\alpha\beta} \lambda x_{o\beta} \lambda z_\alpha \cdot \exists t_\beta \cdot xt \wedge z = ft]$ (image of a set under a function)
$\{\mathbf{F}_\alpha \mid \mathbf{A}_o\}$	stands for $[\lambda \mathbf{z}_\alpha \exists \mathbf{x}^1 \dots \exists \mathbf{x}^n \cdot \mathbf{A}_o \wedge \mathbf{z}_\alpha = \mathbf{F}_\alpha],$ where $\mathbf{x}^1, \dots, \mathbf{x}^n$ are the variables which occur free in both \mathbf{A}_o and \mathbf{F}_α , and \mathbf{z}_α is the first variable which does not occur free in \mathbf{F}_α or in \mathbf{A}_o .

EXERCISES

Prove the following wffs.

X5200. $x_{o\alpha} \cup y_{o\alpha} = \bigcup_{o\alpha(o(o\alpha))} [\lambda v_{o\alpha} \cdot v = x \vee v = y]$

X5201. $x_{o\alpha} \cap y_{o\alpha} = \bigcap_{o\alpha(o(o\alpha))} [\lambda v_{o\alpha} \cdot v = x \vee v = y]$

X5202. $\#f_{\alpha\beta}[x_{o\beta} \cup y_{o\beta}] = \#f_{\alpha\beta}x_{o\beta} \cup \#f_{\alpha\beta}y_{o\beta}$

X5203. $\#f_{\alpha\beta}[x_{o\beta} \cap y_{o\beta}] \subseteq \#f_{\alpha\beta}x_{o\beta} \cap \#f_{\alpha\beta}y_{o\beta}$

X5204. $\#f_{\alpha\beta}[\bigcup_{o\beta(o(o\beta))} w_{o(o\beta)}]$
 $= \bigcup_{o\alpha(o(o\alpha))} [\#_{o(o\alpha)(o(o\beta))((o\alpha)(o\beta))} \#_{o\alpha(o\beta)(\alpha\beta)} f_{\alpha\beta}] w_{o(o\beta)}$

$$\begin{aligned} \mathbf{X5205. } & \# f_{\alpha\beta} [\bigcap_{o\beta(o(o\beta))} w_{o(o\beta)}] \\ & \subseteq \bigcap_{o\alpha(o(o\alpha))} \#_{o(o\alpha)(o(o\beta))((o\alpha)(o\beta))} f_{\alpha\beta} w_{o(o\beta)} \end{aligned}$$

X5206. Derive X5202 from X5204 and X5200.

X5207. Derive X5203 from X5205 and X5201.

X5208. $\exists s_{\alpha} \forall x_i [[sx \vee p_{\alpha}x] \wedge \sim sx \vee q_{\alpha}x] \equiv \forall y_i [py \vee qy]$

X5209. $P_{(o(o\alpha))(o\alpha)}[u_{o\alpha} \cap v_{o\alpha}] = [Pu] \cap [Pv]$

X5210. $Q_{o\alpha\alpha}x_{\alpha} = \{y_{\alpha} \mid y_{\alpha} = x_{\alpha}\}$

X5211. $y_{o\alpha} = \bigcup_{o\alpha(o(o\alpha))} \{Q_{o\alpha\alpha}x_{\alpha} \mid y_{o\alpha}x_{\alpha}\}$

X5212. $\{f_{\alpha\beta}x_{\beta} \mid g_{\alpha\beta}x_{\beta}\} = \# f_{\alpha\beta}g_{\alpha\beta}$

X5213. A wff D is in *β -normal form* iff it has no wf parts of the form $[\lambda x_{\alpha} B_{\beta}]A_{\alpha}$. Show that every wff of Q_0 can be transformed to β -normal form by a sequence of β -contractions and α -conversions.

§53. Equality and Descriptions

5300. $\vdash [x_{\alpha} = y_{\alpha}] \supset h_{\beta\alpha}x_{\alpha} = h_{\beta\alpha}y_{\alpha}$

Proof:

$$(1) \vdash [x_{\alpha} = y_{\alpha}] \supset [h_{\beta\alpha}x_{\alpha} = h_{\beta\alpha}x_{\alpha}] = [h_{\beta\alpha}x_{\alpha} = h_{\beta\alpha}y_{\alpha}]$$

Sub: Axiom 2; λ

$$(2) \vdash [x_{\alpha} = y_{\alpha}] \supset h_{\beta\alpha}x_{\alpha} = h_{\beta\alpha}y_{\alpha}$$

Rule P: .1, 5200

■

5301. $\vdash [x_{\alpha} = y_{\alpha}] \wedge [f_{\beta\alpha} = g_{\beta\alpha}] \supset f_{\beta\alpha}x_{\alpha} = g_{\beta\alpha}y_{\alpha}$

Proof:

$$(1) x_{\alpha} = y_{\alpha}, f_{\beta\alpha} = g_{\beta\alpha} \vdash f_{\beta\alpha}x_{\alpha} = g_{\beta\alpha}y_{\alpha}$$

5201: hypotheses

Then use the Deduction Theorem and Rule P.

■

5302. $\vdash [x_{\alpha} = y_{\alpha}] = [y_{\alpha} = x_{\alpha}]$

Proof:

- | | |
|--|--|
| (.1) $x_\alpha = y_\alpha \vdash y_\alpha = x_\alpha$
(.2) $\vdash [x_\alpha = y_\alpha] \supset \blacksquare y_\alpha = x_\alpha$
(.3) $\vdash [y_\alpha = x_\alpha] \supset \blacksquare x_\alpha = y_\alpha$
(.4) $\vdash [x_\alpha = y_\alpha] = [y_\alpha = x_\alpha]$ | 5201: hypothesis
Deduction Theorem: .1
Sub: .2
Rule P: .2, .3 |
|--|--|

■

5303. $\vdash [x_\alpha = y_\alpha] \supset \blacksquare [x_\alpha = z_\alpha] = [y_\alpha = z_\alpha]$

Proof: Substitute $[\lambda t_\alpha \blacksquare t_\alpha = z_\alpha]$ for $h_{o\alpha}$ in Axiom 2 $^\alpha$, and apply λ -conversion. ■

Note that the transitive law of equality follows from 5303 by Rule P.

DEFINITIONS:

$\sum_{o(o\alpha)}^1$ stands for $[\lambda p_{o\alpha} \exists y_\alpha \blacksquare p_{o\alpha} = Q_{o\alpha\alpha} y_\alpha]$

$\exists_1 x_\alpha A$ stands for $\sum_{o(o\alpha)}^1 [\lambda x_\alpha A]$

(“ $\exists_1 x_\alpha A$ ” means “there is exactly one x_α such that A ”.)

We shall leave the proofs of the next four theorems to the reader.

5304. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \exists y_\alpha \blacksquare p_{o\alpha} = Q_{o\alpha\alpha} y_\alpha$

5305. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \exists y_\alpha \forall z_\alpha \blacksquare p_{o\alpha} z_\alpha = \blacksquare y_\alpha = z_\alpha$

5306. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \exists y_\alpha \blacksquare p_{o\alpha} y_\alpha \wedge \forall z_\alpha \blacksquare p_{o\alpha} z_\alpha \supset y_\alpha = z$

5307. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \blacksquare \exists y_\alpha p_{o\alpha} y_\alpha \wedge \forall y_\alpha \forall z_\alpha \blacksquare p_{o\alpha} y_\alpha \wedge p_{o\alpha} z_\alpha \supset y_\alpha = z_\alpha$

We next introduce description operators for all types.

DEFINITION. $\iota_{o(oo)}$ stands for $Q_{o(oo)(oo)}[\lambda x_o x_o]$

5308. $\vdash \iota_{o(oo)}[Q_{ooo} y_o] = y_o$

Proof:

$$(1) \vdash \iota_{o(oo)}[Q_{ooo} y_o] = \forall x_o \blacksquare [\lambda x_o x_o] x_o = [Q_{ooo} y_o] x_o$$

Sub: Axiom 3, def of ι

(.2)	$\vdash \iota_{o(oo)}[Q_{ooo}y_o] = \forall x_o \bullet x_o = \bullet y_o = x_o$	λ
(.3)	$\vdash [x_o = \bullet y_o = x_o] = y_o$	Rule P
(.4)	$\vdash \iota_{o(oo)}[Q_{ooo}y_o] = \forall x_o y_o$	R: .2, .3
(.5)	$\vdash [\forall x_o y_o] = y_o$	Rule Q
(.6)	$\vdash \iota_{o(oo)}[Q_{ooo}y_o] = y_o$	R: .4, .5

■

REMARK. We could have defined $\iota_{o(oo)}$ as $[\lambda g_{oo} \bullet g_{oo} T_o]$. Then $\iota_{o(oo)}[Q_{ooo}y_o] = y_o$ would have followed immediately from the tautology $[y_o = T] = y_o$. We also could have defined $\iota_{o(oo)}$ as $Q_{o(oo)(oo)}[Q_{ooo}T_o]$.

DEFINITION. We define $\iota_{\gamma(o\gamma)}$ for all type symbols γ by induction on the length of γ and let $[\gamma z_\gamma A]$ stand for $\iota_{\gamma(o\gamma)}[\lambda z_\gamma A]$. Then " $[\gamma z_\gamma A]$ " means "that z_γ such that A ", when there is a unique such z_γ . When $\gamma = (\alpha\beta)$ and $\iota_{\alpha(o\alpha)}$ is already defined, $\iota_{(\alpha\beta)(o(\alpha\beta))}$ stands for

$$[\lambda h_{o(\alpha\beta)} \lambda x_\beta \bullet \gamma z_\alpha \exists f_{\alpha\beta} \bullet h_{o(\alpha\beta)} f_{\alpha\beta} \wedge f_{\alpha\beta} x_\beta = z_\alpha].$$

5309^γ. $\vdash \iota_{\gamma(o\gamma)}[Q_{o\gamma\gamma}y_\gamma] = y_\gamma$ for each type symbol γ .

Proof: We prove this by induction on the length of γ . 5309[°] is Axiom 5, and 5309[°] is 5308. Suppose $\gamma = (\alpha\beta)$.

(.1)	$\vdash \iota_{(\alpha\beta)(o(\alpha\beta))}[Q_{o(\alpha\beta)(\alpha\beta)}y_{\alpha\beta}] = [\lambda x_\beta \gamma z_\alpha \exists f_{\alpha\beta} \bullet y_{\alpha\beta} = f_{\alpha\beta} \wedge f_{\alpha\beta} x_\beta = z_\alpha]$	by λ and the definitions of $=$ and $\iota_{\gamma(o\gamma)}$
(.2)	$y_{\alpha\beta} = f_{\alpha\beta} \wedge f_{\alpha\beta} x_\beta = z_\alpha \vdash y_{\alpha\beta} x_\beta = z_\alpha$	Rule P, $=$, Rule R': hyp
(.3)	$\vdash \exists f_{\alpha\beta}[y_{\alpha\beta} = f_{\alpha\beta} \wedge f_{\alpha\beta} x_\beta = z_\alpha] \supset y_{\alpha\beta} x_\beta = z_\alpha$	

\exists Rule (5244), Deduction Theorem: .2

(.4)	$y_{\alpha\beta} x_\beta = z_\alpha \vdash y_{\alpha\beta} = y_{\alpha\beta} \wedge y_{\alpha\beta} x_\beta = z_\alpha$	Rule P: 5200, hyp
(.5)	$y_{\alpha\beta} x_\beta = z_\alpha \vdash \exists f_{\alpha\beta} \bullet y_{\alpha\beta} = f_{\alpha\beta} \wedge f_{\alpha\beta} x_\beta = z_\alpha$	\exists Gen: .4
(.6)	$\vdash \exists f_{\alpha\beta}[y_{\alpha\beta} = f_{\alpha\beta} \wedge f_{\alpha\beta} x_\beta = z_\alpha] = \bullet y_{\alpha\beta} x_\beta = z_\alpha$	

Deduction Theorem, Rule P: 3, .5

(.7)	$\vdash \iota_{(\alpha\beta)(o(\alpha\beta))}[Q_{o(\alpha\beta)(\alpha\beta)}y_{\alpha\beta}] = [\lambda x_\beta \bullet \iota_{\alpha(o\alpha)} \lambda z_\alpha \bullet Q_{o\alpha\alpha}[y_{\alpha\beta} x_\beta] z_\alpha]$	R: .1, .6, def of $=$
------	---	-----------------------

(.8)	$\vdash \iota_{(\alpha\beta)(o(\alpha\beta))}[Q_{o(\alpha\beta)(\alpha\beta)}y_{\alpha\beta}] = [\lambda x_\beta \bullet \iota_{\alpha(o\alpha)} Q_{o\alpha\alpha}[y_{\alpha\beta} x_\beta]]$	Sub: 5205; Rule RR: .7
------	---	------------------------

(.9)	$\vdash \iota_{(\alpha\beta)(o(\alpha\beta))}[Q_{o(\alpha\beta)(\alpha\beta)}y_{\alpha\beta}] = [\lambda x_\beta \bullet y_{\alpha\beta} x_\beta]$	Sub: 5309 ^α (induction hypothesis); Rule R: .8
------	--	---

(.10)	$\vdash \iota_{(\alpha\beta)(o(\alpha\beta))}[Q_{o(\alpha\beta)(\alpha\beta)}y_{\alpha\beta}] = y_{\alpha\beta}$	RR: 5205, .9
-------	--	--------------

■

5310. $\vdash \forall z_\alpha [p_{o\alpha} z_\alpha = \blacksquare y_\alpha = z_\alpha] \supset \blacksquare \iota_{\alpha(o\alpha)} p_{o\alpha} = y_\alpha$

Proof:

- (.1) $.1 \vdash \forall z_\alpha \blacksquare p_{o\alpha} z_\alpha = [Q_{o\alpha\alpha} y_\alpha] z_\alpha$ hyp
- (.2) $.1 \vdash p_{o\alpha} = Q_{o\alpha\alpha} y_\alpha$ Sub: Axiom 3; RR: .1
- (.3) $.1 \vdash \iota_{\alpha(o\alpha)} p_{o\alpha} = y_\alpha$ $=: .2, R': 5309$

Then use the Deduction Theorem. ■

5311. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha \supset p_{o\alpha} [\iota_{\alpha(o\alpha)} p_{o\alpha}]$

Proof:

- (.1) $.1 \vdash p_{o\alpha} = \blacksquare Q_{o\alpha\alpha} y_\alpha$ hyp
- (.2) $.1 \vdash p_{o\alpha} y_\alpha = \blacksquare Q_{o\alpha\alpha} y_\alpha y_\alpha$ $=: .1$
- (.3) $.1 \vdash p_{o\alpha} y_\alpha$ $=, R': .2, 5200$
- (.4) $.1 \vdash p_{o\alpha} [\iota_{\alpha(o\alpha)} \blacksquare Q_{o\alpha\alpha} y_\alpha]$ RR: 5309, .3
- (.5) $.1 \vdash p_{o\alpha} [\iota_{\alpha(o\alpha)} p_{o\alpha}]$ $=, R': .1, .4$
- (.6) $\vdash \exists y_\alpha [p_{o\alpha} = \blacksquare Q_{o\alpha\alpha} y_\alpha] \supset p_{o\alpha} [\iota_{\alpha(o\alpha)} p_{o\alpha}]$

\exists Rule, Deduction Theorem: .5

Then use 5304 and .6 to obtain the desired theorem. ■

5312. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha \supset \forall z_\alpha \blacksquare p_{o\alpha} z_\alpha = \blacksquare \iota_{\alpha(o\alpha)} p_{o\alpha} = z_\alpha$

Proof:

- (.1) $.1 \vdash p_{o\alpha} = Q_{o\alpha\alpha} y_\alpha$ hyp
- (.2) $.1 \vdash \iota_{\alpha(o\alpha)} p_{o\alpha} = y_\alpha$ $=: .1, 5309$
- (.3) $.1 \vdash p_{o\alpha} = Q_{o\alpha\alpha} [\iota_{\alpha(o\alpha)} p_{o\alpha}]$ $=, R': .1, .2$
- (.4) $.1 \vdash \forall z_\alpha \blacksquare p_{o\alpha} z_\alpha = \blacksquare \iota_{\alpha(o\alpha)} p_{o\alpha} = z_\alpha$ Sub: Axiom 3; R': .3
- (.5) $\vdash \exists y_\alpha [p_{o\alpha} = Q_{o\alpha\alpha} y_\alpha] \supset \forall z_\alpha \blacksquare p_{o\alpha} z_\alpha = \blacksquare \iota_{\alpha(o\alpha)} p_{o\alpha} = z_\alpha$

\exists Rule, Deduction Theorem: .4

The Theorem follows from .5 and 5304. ■

DEFINITION: Let $C_{\gamma o\gamma\gamma}$ be

$$[\lambda x_\gamma \lambda y_\gamma \lambda p_o \lambda q_\gamma \blacksquare [p_o \wedge \blacksquare x_\gamma = q_\gamma] \vee [\sim p_o \wedge \blacksquare y_\gamma = q_\gamma]]$$

$C_{\gamma o\gamma\gamma} x_\gamma y_\gamma p_o$ can be read “if p_o then x_γ , else y_γ ”.

5313. $\vdash [C_{\gamma o\gamma\gamma} x_\gamma y_\gamma T_o = x_\gamma] \wedge \blacksquare C_{\gamma o\gamma\gamma} x_\gamma y_\gamma F_o = y_\gamma$

Proof:

- (.1) $\vdash C_{\gamma\alpha\gamma\gamma}x_\gamma y_\gamma T = \exists q_\gamma \bullet [T_o \wedge \bullet x_\gamma = q_\gamma] \vee \bullet \sim T \wedge \bullet y_\gamma = q_\gamma \quad \lambda$
- (.2) $\vdash [[T \wedge \bullet x_\gamma = q_\gamma] \vee \bullet \sim T \wedge \bullet y_\gamma = q_\gamma] = \bullet x_\gamma = q_\gamma \quad \text{Rule P}$
- (.3) $\vdash C_{\gamma\alpha\gamma\gamma}x_\gamma y_\gamma T = \iota_{\gamma(\alpha)} \lambda q_\gamma [x_\gamma = q_\gamma] \quad \text{R: .1, .2}$
- (.4) $\vdash C_{\gamma\alpha\gamma\gamma}x_\gamma y_\gamma T = \iota_{\gamma(\alpha)} [Q_{\alpha\gamma\gamma}x_\gamma] \quad \text{Sub: 5205; RR: .3; def of =}$
- (.5) $\vdash C_{\gamma\alpha\gamma\gamma}x_\gamma y_\gamma T = x_\gamma \quad \text{Sub: 5309; R: .4}$
- (.6) $\vdash C_{\gamma\alpha\gamma\gamma}x_\gamma y_\gamma F = \iota_{\gamma(\alpha)} \lambda q_\gamma [y_\gamma = q_\gamma] \quad \lambda, \text{ Rule P, R}$
- (.7) $\vdash C_{\gamma\alpha\gamma\gamma}x_\gamma y_\gamma F = y_\gamma \quad \text{Sub: 5205; RR: .6, 5309}$

Then use Rule P with .5 and .7. ■

DEFINITION: Let AC^α be $\exists j_{\alpha(o\alpha)} \forall p_{o\alpha} \bullet \exists x_\alpha p_{o\alpha} x_\alpha \supset p_{o\alpha} \bullet j_{\alpha(o\alpha)} p_{o\alpha}$.

This is a formulation of the Axiom of Choice. AC^α says that there is a choice function j which chooses from every nonempty set p (of type $(o\alpha)$) an element (which is designated as jp) of that set. Normally when one assumes the Axiom of Choice in type theory, one assumes it as an axiom schema, and asserts AC^α for each type symbol α . For more information about the Axiom of Choice see [Moore, 1982] and [Rubin and Rubin, 1985].

EXERCISES

X5300. Prove $\sum_{o(o\alpha)}^1 p_{o\alpha} = \bullet p_{o\alpha} = Q_{o\alpha\alpha} \bullet \iota_{\alpha(o\alpha)} p_{o\alpha}$

X5301. Prove that for each type symbol α ,

$$\forall x_i \forall y_i [x_i = y_i] \supset AC^\alpha.$$

X5302. Prove that if β occurs in α , then $\vdash AC^\alpha \supset AC^\beta$. (*Hint:* X5301 may be useful.)

X5303. Prove $Q_{o\alpha\alpha} = [\lambda x_\alpha \lambda y_\alpha \forall p_{o\alpha\alpha} \bullet \forall z_\alpha p_{o\alpha\alpha} z_\alpha z_\alpha \supset p_{o\alpha\alpha} x_\alpha y_\alpha]$

X5304. A famous theorem due to Georg Cantor says that the power set (set of all subsets) of a set always has greater cardinality than the set. This simply means that every set has more subsets than members, and this can be expressed simply by saying that there is no function which maps a set onto its power set. The latter statement can be expressed in type theory by the following wff: $\forall s_{o\alpha} \sim \exists g_{o\alpha\alpha} \forall f_{o\alpha} \bullet [f_{o\alpha} \subseteq s_{o\alpha}] \supset \exists j_\alpha \bullet s_{o\alpha} j_\alpha \wedge g_{o\alpha\alpha} j_\alpha = f_{o\alpha}$. (Here α is an arbitrary type symbol.) If we take $s_{o\alpha}$ to be $[\lambda x_\alpha T_o]$, the wff above reduces to

$$\sim \exists g_{o\alpha\alpha} \forall f_{o\alpha} \exists j_\alpha \bullet g_{o\alpha\alpha} j_\alpha = f_{o\alpha}$$

Prove this wff in \mathcal{Q}_0 .

X5305. Prove $\forall s_{o\alpha} \sim \exists g_{o\alpha\alpha} \forall f_{o\alpha} \bullet [f_{o\alpha} \subseteq s_{o\alpha}] \supset \exists j_\alpha \bullet s_{o\alpha} j_\alpha \wedge g_{o\alpha\alpha} j_\alpha = f_{o\alpha}$

X5306. A variant of the second wff in exercise X5304 is

$$\exists s_n \forall m_i \sim [s_n m_i = m_i] \supset \sim \exists g_{nn} \forall f_n \exists j_i \bullet g_{nn} j_i = f_n$$

Explain what this says intuitively, and prove it.

X5307. Let α and β be arbitrary type symbols, and let $P^{\alpha\beta}$ (pairing function) be the wff

$$[\lambda x_\alpha \lambda y_\beta \lambda g_{o\beta\alpha} \bullet g_{o\beta\alpha} x_\alpha y_\beta].$$

(a) What is the type of the wff $P^{\alpha\beta}$? Let $L^{\alpha\beta}$ (left projection) be the wff

$$[\lambda p_\delta \bullet \forall x_\alpha \exists y_\beta \bullet p_\delta = P^\alpha x_\alpha y_\beta],$$

where δ is chosen so that $L^{\alpha\beta}$ is a wff.

- (b) What type symbol must δ be?
 (c) Prove the following theorem of \mathcal{Q}_0 :

$$L^{\alpha\beta}[P^{\alpha\beta} x_\alpha y_\beta] = x_\alpha.$$

(d) Define a wff $R^{\alpha\beta}$ (right projection) such that $[R^{\alpha\beta}[P^{\alpha\beta} x_\alpha y_\beta] = y_\beta]$ is a theorem \mathcal{Q}_0 . What is the type of $R^{\alpha\beta}$?

(e) To define the left projection operator another way, define $M^{\alpha\beta}$ as
 $[\lambda p_{o(o\beta\alpha)} \bullet \forall x_\alpha \exists y_\beta \bullet g_{o\beta\alpha} \bullet p_{o(o\beta\alpha)} g_{o\beta\alpha} \supset \exists y_\beta \bullet g_{o\beta\alpha} x_\alpha y_\beta].$
 Prove $M^{\alpha\beta}[P^{\alpha\beta} x_\alpha y_\beta] = x_\alpha$.

X5308. Prove $AC^\beta \supset \bullet \forall x_\alpha \exists y_\beta p_{o\beta\alpha} x_\alpha y_\beta \equiv \exists f_{\beta\alpha} \forall x_\alpha \bullet p_{o\beta\alpha} x_\alpha \bullet f_{\beta\alpha} x_\alpha$

X5309. Prove $\sim \exists h_{\iota(o\iota)} \forall p_{o\iota} \forall q_{o\iota} \cdot h p = h q \supset p = q$

This is called the Injective Cantor Theorem, since it asserts that there is no injection from the power set of the domain of individuals into that domain.

X5310. Prove $\forall r_{o\beta(o\beta)} [\forall x_{o\beta} \exists y_\beta r x y \supset \exists f_{\beta(o\beta)} \forall x r x. f x] \supset AC^\beta$
 (Compare this with X5308.)

X5311. Prove $\sim [x_\iota = y_\iota] \supset \exists f_\iota \bullet f x = a_\iota \wedge f y = b_\iota$

§54. Semantics of \mathcal{Q}_0

A *frame* is a collection $\{\mathcal{D}_\alpha\}_\alpha$ of nonempty domains (sets) \mathcal{D}_α , one for each type symbol α , such that $\mathcal{D}_o = \{\top, \perp\}$ and $\mathcal{D}_{\alpha\beta}$ is some collection of functions mapping \mathcal{D}_β into \mathcal{D}_α . The members of \mathcal{D}_o are called *truth values* and the members of \mathcal{D}_i are called *individuals*.

If $x \in \mathcal{D}_\alpha$, the one-element set whose sole member is x (sometimes denoted $\{x\}$) is that function h from \mathcal{D}_α into \mathcal{D}_o such that $hx = \top$ and $hy = \perp$ for each element y of \mathcal{D}_α distinct from x . The *identity relation* on \mathcal{D}_α is that function q with domain \mathcal{D}_α such that for each $x \in \mathcal{D}_\alpha$, $qx = \{x\}$. Thus if x and y are in \mathcal{D}_α , $(qx)y = \top$ iff $x = y$.

An *interpretation* $\langle\{\mathcal{D}_\alpha\}_\alpha, \mathcal{J}\rangle$ of \mathcal{Q}_0 consists of a frame and a function \mathcal{J} which maps each constant of type α of \mathcal{Q}_0 to an element of \mathcal{D}_α , so that $\mathcal{J}Q_{o\alpha\alpha}$ is the identity relation on \mathcal{D}_α , and $\mathcal{J}I_{i(\alpha)}$ is some function from \mathcal{D}_o into \mathcal{D}_i which maps each one-element set in \mathcal{D}_o to its unique member in \mathcal{D}_i . For any constant C , $\mathcal{J}C$ is called the *denotation* of C .

An *assignment* into a frame $\{\mathcal{D}_\alpha\}_\alpha$ (sometimes called an assignment of values in the frame $\{\mathcal{D}_\alpha\}_\alpha$ to the variables of \mathcal{Q}_0) is a function φ with domain the set of variables of \mathcal{Q}_0 such that for each variable x_α , $\varphi x_\alpha \in \mathcal{D}_\alpha$. An assignment into an interpretation is an assignment into the frame of the interpretation. In contexts where an interpretation is under discussion, it will be assumed that all assignments discussed are into that interpretation unless otherwise indicated. Given an assignment φ , a variable x_α , and $z \in \mathcal{D}_\alpha$, let $(\varphi : x_\alpha/z)$ be that assignment ψ such that $\psi x_\alpha = z$ and $\psi y_\beta = \varphi y_\beta$ if $y_\beta \neq x_\alpha$.

An interpretation $\mathcal{M} = \langle\{\mathcal{D}_\alpha\}_\alpha, \mathcal{J}\rangle$ is a *general model* for \mathcal{Q}_0 iff there is a binary function $\mathcal{V}^{\mathcal{M}}$ such that for each assignment φ and wff A_α , $\mathcal{V}_\varphi^{\mathcal{M}} A_\alpha \in \mathcal{D}_\alpha$, and the following conditions are satisfied for all assignments φ and all wffs:

- (a) $\mathcal{V}_\varphi^{\mathcal{M}} x_\alpha = \varphi x_\alpha$.
- (b) $\mathcal{V}_\varphi^{\mathcal{M}} A_\alpha = \mathcal{J}A_\alpha$ if A_α is a primitive constant.
- (c) $\mathcal{V}_\varphi^{\mathcal{M}} [A_{\alpha\beta} B_\beta] = (\mathcal{V}_\varphi^{\mathcal{M}} A_{\alpha\beta})(\mathcal{V}_\varphi^{\mathcal{M}} B_\beta)$ (the value of a function $\mathcal{V}_\varphi^{\mathcal{M}} A_{\alpha\beta}$ at the argument $\mathcal{V}_\varphi^{\mathcal{M}} B_\beta$).
- (d) $\mathcal{V}_\varphi^{\mathcal{M}} [\lambda x_\alpha B_\beta] =$ that function from \mathcal{D}_α into \mathcal{D}_β whose value for each argument $z \in \mathcal{D}_\alpha$ is $\mathcal{V}_{(\varphi : x_\alpha/z)}^{\mathcal{M}} B_\beta$.

If an interpretation \mathcal{M} is a general model, the function $\mathcal{V}^{\mathcal{M}}$ is uniquely determined. $\mathcal{V}_\varphi^{\mathcal{M}} A_\alpha$ is called the *value* of A_α in \mathcal{M} with respect to φ . We

sometimes write $\mathcal{V}_{\varphi}^{\mathcal{M}}$ as \mathcal{V}_{φ} , as $\mathcal{V}^{\mathcal{M}}$, or as \mathcal{V} when the interpretation or assignment is clear from the context, or irrelevant.

REMARK. Not all frames belong to interpretations, and not all interpretations are general models. In order to be a general model, an interpretation must have a frame satisfying certain closure conditions which are discussed further in [Andrews, 1972b]. Basically, in a general model every wff must have a value with respect to each assignment. Thus, the identity function mapping \mathcal{D}_i onto itself must be a member of \mathcal{D}_{ii} so that $\mathcal{V}_{\varphi}^{\mathcal{M}}[\lambda x_i x_i]$ will be defined. Similarly, if \mathcal{M} is a model where \mathcal{D}_i is the set of natural numbers and \mathcal{D}_{ii} contains the addition function j such that $jxy = x + y$ for all numbers x and y , then \mathcal{D}_{ii} must contain the function k such that $kx = 2x + 5$ for all numbers x , since $k = \mathcal{V}_{\varphi}^{\mathcal{M}}[\lambda x_i \cdot f_{ii}[f_{ii}x_i x_i]y_i]$ when $\varphi f_{ii} = j$ and $\varphi y_i = 5$.

An interpretation $\langle \{\mathcal{D}_{\alpha}\}_{\alpha}, \mathcal{J} \rangle$ is a *standard model* iff for all α and β , $\mathcal{D}_{\alpha\beta}$ is the set of all functions from \mathcal{D}_{β} into \mathcal{D}_{α} . Clearly a standard model is a general model.

5400 Proposition. Let \mathcal{M} be a general model, \mathbf{A}_{α} a wff, and φ and ψ assignments which agree on all free variables of \mathbf{A}_{α} . Then $\mathcal{V}_{\varphi}^{\mathcal{M}}\mathbf{A}_{\alpha} = \mathcal{V}_{\psi}^{\mathcal{M}}\mathbf{A}_{\alpha}$.

Proof: By induction on the construction of \mathbf{A}_{α} . The details are left to the reader. ■

It is clear that if \mathbf{A}_{α} is a closed wff, then $\mathcal{V}^{\mathcal{M}}\mathbf{A}_{\alpha}$ may be considered meaningful without regard to any assignment. In this case, $\mathcal{V}^{\mathcal{M}}\mathbf{A}_{\alpha}$ is called the *denotation* of \mathbf{A}_{α} in \mathcal{M} .

DEFINITIONS: Let \mathbf{A} be a wff_o, \mathcal{M} a general model, and φ an assignment into \mathcal{M} .

- (1) φ satisfies \mathbf{A} in \mathcal{M} iff $\mathcal{V}_{\varphi}^{\mathcal{M}}\mathbf{A} = \top$.
- (2) \mathbf{A} is *satisfiable* in \mathcal{M} iff there is an assignment which satisfies \mathbf{A} in \mathcal{M} .
- (3) \mathbf{A} is *valid* in \mathcal{M} iff for every assignment φ into \mathcal{M} , $\mathcal{V}_{\varphi}^{\mathcal{M}}\mathbf{A} = \top$.
- (4) A sentence \mathbf{A} is *true* in \mathcal{M} iff $\mathcal{V}^{\mathcal{M}}\mathbf{A} = \top$, and *false* in \mathcal{M} iff $\mathcal{V}^{\mathcal{M}}\mathbf{A} = \perp$.
- (5) \mathbf{A} is *valid in the general [standard] sense* iff \mathbf{A} is valid in every general [standard] model of \mathcal{Q}_0 .
- (6) A *model for a set \mathcal{G} of wffs_o* is a model for \mathcal{Q}_0 in which each wff of \mathcal{G} is valid.

We write $\mathcal{M} \models_{\varphi} \mathbf{A}$, $\mathcal{M} \models \mathbf{A}$, and $\models \mathbf{A}$ to indicate that φ satisfies \mathbf{A} in \mathcal{M} , \mathbf{A} is valid in \mathcal{M} , and \mathbf{A} is valid in the general sense, respectively.

Clearly a wff which is valid in the general sense is valid in the standard sense, though we shall see in Chapter 7 that the converse of this statement is false. We shall show that a wff_o of \mathcal{Q}_0 is a theorem iff it is valid in the general sense.

5401 Lemma. Let \mathcal{M} be a general model, and φ be an assignment into \mathcal{M} .

- (a) $\mathcal{V}_{\varphi}^{\mathcal{M}}[[\lambda x_{\alpha} B_{\beta}] A_{\alpha}] = \mathcal{V}_{(\varphi : x_{\alpha} / \mathcal{V}_{\varphi}^{\mathcal{M}} A_{\alpha})}^{\mathcal{M}} B_{\beta}$.
- (b) $\mathcal{V}_{\varphi}^{\mathcal{M}}[A_{\alpha} = B_{\alpha}] = T$ iff $\mathcal{V}_{\varphi}^{\mathcal{M}} A_{\alpha} = \mathcal{V}_{\varphi}^{\mathcal{M}} B_{\alpha}$.
- (c) $\mathcal{V}^{\mathcal{M}} T_o = T$.
- (d) $\mathcal{V}^{\mathcal{M}} F_o = F$.
- (e) If $x, y \in \mathcal{D}_o$, then $(\mathcal{V}^{\mathcal{M}} \wedge_{ooo})xy = T$ if $x = T$ and $y = T$;
 $= F$ otherwise.
- (f) If $x, y \in \mathcal{D}_o$, then $(\mathcal{V}^{\mathcal{M}} \supset_{ooo})xy = T$ if $x = F$ or $y = T$;
 $= F$ if $x = T$ and $y = F$.
- (g) $\mathcal{M} \models_{\varphi} \forall x A$ iff $\mathcal{M} \models_{\psi} A$ for all assignments ψ which agree with φ off x_{α} .

Proof: These results follow rather directly from the relevant definitions, so we leave most of the details to the reader. In the case of (e), one shows the aid of (a) and (b) that if $\varphi x_o = x$ and $\varphi y_o = y$, then $(\mathcal{V}^{\mathcal{M}} \wedge_{ooo})xy = T$ iff $\mathcal{V}[\lambda g_{ooo} \bullet g_{ooo}TT] = \mathcal{V}_{\varphi}[\lambda g_{ooo} \bullet g_{ooo}x_oy_o]$, which is true iff for every function $g \in \mathcal{D}_{ooo}$, $gTT = (\mathcal{V}[\lambda g_{ooo} \bullet g_{ooo}TT])g = (\mathcal{V}_{\varphi}[\lambda g_{ooo} \bullet g_{ooo}x_oy_o])g = gxy$. Clearly $gTT = gxy$ for every $g \in \mathcal{D}_{ooo}$ if $x = T = y$. On the other hand, if $x \neq T$ (for example), then $(\mathcal{V}[\lambda x_o \lambda y_o x_o])TT = T \neq x = \mathcal{V}_{\varphi}x_o = (\mathcal{V}_{\varphi}[\lambda y_o x_o])y = (\mathcal{V}[\lambda x_o \lambda y_o x_o])xy$, so $(\mathcal{V}^{\mathcal{M}} \wedge_{ooo})xy = F$.

In the case of (g), one sees that $\models_{\varphi} \forall x_{\alpha} A$ iff $\mathcal{V}_{\varphi}[\lambda x_{\alpha} T] = \mathcal{V}_{\varphi}[\lambda x_{\alpha} A]$, which is true iff for every $z \in \mathcal{D}_{\alpha}$, $T = (\mathcal{V}_{\varphi}[\lambda x_{\alpha} T])z = (\mathcal{V}_{\varphi}[\lambda x_{\alpha} A])z = \mathcal{V}_{(\varphi : x_{\alpha} / z)} A$. This is equivalent to the stated condition. ■

REMARK. Speaking loosely, we might expect that a sentence $\forall x_{\alpha\beta} A$ is true in a model $\mathcal{M} = \langle \{\mathcal{D}_{\alpha}\}_{\alpha}, \mathcal{I} \rangle$ iff A is true for all subsets x of \mathcal{D}_{β} . However, we see from 5401(g) that it would be more accurate to say that $\forall x_{\alpha\beta} A$ is true in \mathcal{M} iff A is true for all of those subsets x of \mathcal{D}_{β} which are in $\mathcal{D}_{o\beta}$. When \mathcal{M} is nonstandard, these are quite different statements.

5402 Soundness Theorem.

- (a) Every theorem of \mathcal{Q}_0 is valid in the general sense, and hence in the standard sense.

(b) If \mathcal{G} is a set of wffs _{σ} and \mathcal{M} is a model for \mathcal{G} and $\mathcal{G} \vdash \mathbf{A}$, then $\mathcal{M} \models \mathbf{A}$.

Proof: We show that Rule R preserves validity in every general model for \mathcal{Q}_0 , and that the axioms of \mathcal{Q}_0 are valid in the general sense.

Rule R. Suppose that \mathcal{M} is a general model such that $\mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{A}_{\alpha} = \mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{B}_{\alpha}$ for all assignments φ into \mathcal{M} , and \mathbf{C}_{β} and \mathbf{C}'_{β} are wffs such that \mathbf{C}'_{β} is obtained from \mathbf{C}_{β} by replacing at most one occurrence of \mathbf{A}_{α} (which is not a variable immediately preceded by λ) in \mathbf{C}_{β} by an occurrence of \mathbf{B}_{α} . We prove that $\mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{C}_{\beta} = \mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{C}'_{\beta}$ for all assignments φ , by induction on the construction of \mathbf{C}_{β} .

Case: \mathbf{C}'_{β} is \mathbf{C}_{β} . This case is trivial.

Case: \mathbf{C}_{β} is \mathbf{A}_{α} . Then \mathbf{C}'_{β} is \mathbf{B}_{α} or \mathbf{A}_{α} , so this case is trivial.

In the remaining cases we assume that \mathbf{A}_{α} is a proper part of \mathbf{C}_{β} .

Case: \mathbf{C}_{β} has the form $[\mathbf{G}_{\beta\gamma} \mathbf{H}_{\gamma}]$.

\mathbf{C}'_{β} has the form $[\mathbf{G}'_{\beta\gamma} \mathbf{H}'_{\gamma}]$, so by inductive hypothesis we have $\mathcal{V}_{\varphi} \mathbf{G}_{\beta\gamma} = \mathcal{V}_{\varphi} \mathbf{G}'_{\beta\gamma}$ and $\mathcal{V}_{\varphi} \mathbf{H}_{\gamma} = \mathcal{V}_{\varphi} \mathbf{H}'_{\gamma}$, so $\mathcal{V}_{\varphi} \mathbf{C}_{\beta} = (\mathcal{V}_{\varphi} \mathbf{G}_{\beta\gamma})(\mathcal{V}_{\varphi} \mathbf{H}_{\gamma}) = (\mathcal{V}_{\varphi} \mathbf{G}'_{\beta\gamma})(\mathcal{V}_{\varphi} \mathbf{H}'_{\gamma}) = \mathcal{V}_{\varphi} \mathbf{C}'_{\beta}$.

Case: \mathbf{C}_{β} has the form $[\lambda \mathbf{x}_{\gamma} \mathbf{E}_{\delta}]$.

\mathbf{C}'_{β} has the form $[\lambda \mathbf{x}_{\gamma} \mathbf{E}'_{\delta}]$ and by induction hypothesis $\mathcal{V}_{\psi} \mathbf{E}_{\delta} = \mathcal{V}_{\psi} \mathbf{E}'_{\delta}$ for all assignments ψ , so by the definition of \mathcal{V}_{φ} we see that $\mathcal{V}_{\varphi} \mathbf{C}_{\beta} = \mathcal{V}_{\varphi} \mathbf{C}'_{\beta}$.

Now suppose that $\mathcal{M} \models [\mathbf{A}_{\alpha} = \mathbf{B}_{\alpha}]$ and $\mathcal{M} \models \mathbf{C}$, and \mathbf{C}' is obtained from \mathbf{C} by Rule R. Then $\mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{A}_{\alpha} = \mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{B}_{\alpha}$ for all assignments φ , so for any assignment φ we have $\mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{C}' = \mathcal{V}_{\varphi}^{\mathcal{M}} \mathbf{C} = \top$, so $\mathcal{M} \models \mathbf{C}'$.

To show that the axioms are valid, we let \mathcal{M} be any general model and φ be any assignment, and show for each axiom \mathbf{A} that $\mathcal{M} \models_{\varphi} \mathbf{A}$.

Axiom 1. Suppose first that $(\varphi g_{oo})\top = \top$ and $(\varphi g_{oo})F = \top$. Then by Lemma 5401 $\mathcal{V}_{\varphi}[g_{oo}\top \wedge g_{oo}F] = \top$, and for any assignment ψ which agrees with φ off x_o , $\mathcal{V}_{\psi}[g_{oo}x_o] = (\psi g_{oo})(\psi x_o) = (\varphi g_{oo})(\psi x_o) = \top$, so $\mathcal{V}_{\varphi}\forall x_o[g_{oo}x_o] = \top$ by 5401(g). Hence $\models_{\varphi}[\text{Axiom 1}]$ in this case.

On the other hand suppose there exists $z \in \mathcal{D}_o$ such that $(\varphi g_{oo})z = F$. Define ψ to agree with φ off x_o , with $\psi x_o = z$. Then $\mathcal{V}_{\psi}[g_{oo}x_o] = (\psi g_{oo})(\psi x_o) = (\varphi g_{oo})z = F$, so $\mathcal{V}_{\varphi}\forall x_o[g_{oo}x_o] = F$ by 5401(g). But $\mathcal{V}_{\varphi}[g_{oo}\top \wedge g_{oo}F] = F$ too, so $\models_{\varphi}[\text{Axiom 1}]$ in this case also.

Axiom 2. If $\varphi x_\alpha \neq \varphi y_\alpha$, then \models_φ [Axiom 2] by 5401(b) and (f). If $\varphi x_\alpha = \varphi y_\alpha$, then $\mathcal{V}_\varphi[h_{o\alpha}x_\alpha] = (\varphi h_{o\alpha})(\varphi x_\alpha) = (\varphi h_{o\alpha})(\varphi y_\alpha) = \mathcal{V}_\varphi[h_{o\alpha}y_\alpha]$, so \models_φ [Axiom 2] by 5401(b) and (f).

Axiom 3. Suppose $\varphi f_{\alpha\beta} = \varphi g_{\alpha\beta}$, and let ψ be any assignment which agrees with φ off x_β . Then $\mathcal{V}_\psi[f_{\alpha\beta}x_\beta] = (\psi f_{\alpha\beta})(\psi x_\beta) = (\varphi f_{\alpha\beta})(\psi x_\beta) = (\varphi g_{\alpha\beta})(\psi x_\beta) = (\psi g_{\alpha\beta})(\psi x_\beta) = \mathcal{V}_\psi[g_{\alpha\beta}x_\beta]$, so $\mathcal{V}_\psi[f_{\alpha\beta}x_\beta = g_{\alpha\beta}x_\beta] = \top$, so $\mathcal{V}_\varphi \forall x_\beta [f_{\alpha\beta}x_\beta = g_{\alpha\beta}x_\beta] = \top$. Also, $\mathcal{V}_\varphi[f_{\alpha\beta} = g_{\alpha\beta}] = \top$, so \models_φ [Axiom 3].

Suppose $(\varphi f_{\alpha\beta}) \neq (\varphi g_{\alpha\beta})$. Then there is an element $z \in \mathcal{D}_\beta$ such that $(\varphi f_{\alpha\beta})z \neq (\varphi g_{\alpha\beta})z$. Let $\psi = (\varphi : x_\beta/z)$. Then $\mathcal{V}_\psi[f_{\alpha\beta}x_\beta] = (\psi f_{\alpha\beta})(\psi x_\beta) = (\varphi f_{\alpha\beta})z \neq (\varphi g_{\alpha\beta})z = (\psi g_{\alpha\beta})(\psi x_\beta) = \mathcal{V}_\psi[g_{\alpha\beta}x_\beta]$, so $\mathcal{V}_\psi[f_{\alpha\beta}x_\beta = g_{\alpha\beta}x_\beta] = \text{F}$, and $\mathcal{V}_\varphi \forall x_\beta [f_{\alpha\beta}x_\beta = g_{\alpha\beta}x_\beta] = \text{F}$ by 5401(g). Also $\mathcal{V}_\varphi[f_{\alpha\beta} = g_{\alpha\beta}] = \text{F}$, so \models_φ [Axiom 3].

For each of the Axioms 4₁ – 4₅, the left side of the equality has the form $[[\lambda x_\alpha D_\beta] A_\alpha]$. Let $\psi = (\varphi : x_\alpha/\mathcal{V}_\varphi A_\alpha)$. Then by 5401(a) $\mathcal{V}_\varphi[[\lambda x_\alpha D_\beta] A_\alpha] = \mathcal{V}_\psi D_\beta$.

Axiom 4₁. $[\lambda x_\alpha D_\beta] A_\alpha = D_\beta$, where D_β is a primitive constant or variable distinct from x_α . $\mathcal{V}_\psi D_\beta = \mathcal{V}_\varphi D_\beta$ by 5400, so \models_φ [Axiom 4₁].

Axiom 4₂. $[\lambda x_\alpha x_\alpha] A_\alpha = A_\alpha$. Here D_β is x_α and $\mathcal{V}_\psi D_\beta = \psi x_\alpha = \mathcal{V}_\varphi A_\alpha$, so \models_φ [Axiom 4₂].

Axiom 4₃. $[\lambda x_\alpha \bullet B_{\beta\gamma} C_\gamma] A_\alpha = [[\lambda x_\alpha B_{\beta\gamma}] A_\alpha][[\lambda x_\alpha C_\gamma] A_\alpha]$. D_β is $[B_{\beta\gamma} C_\gamma]$.

$$\begin{aligned}\mathcal{V}_\psi D_\beta &= (\mathcal{V}_\psi B_{\beta\gamma})(\mathcal{V}_\psi C_\gamma) = (\mathcal{V}_\varphi[[\lambda x_\alpha B_{\beta\gamma}] A_\alpha])(\mathcal{V}_\varphi[[\lambda x_\alpha C_\gamma] A_\alpha]) \\ &= \mathcal{V}_\varphi[[\lambda x_\alpha B_{\beta\gamma}] A_\alpha][[\lambda x_\alpha C_\gamma] A_\alpha],\end{aligned}$$

so \models_φ [Axiom 4₃].

Axiom 4₄. $[\lambda x_\alpha \bullet \lambda y_\gamma B_\delta] A_\alpha = [\lambda y_\gamma \bullet [\lambda x_\alpha B_\delta] A_\alpha]$, where y_γ is distinct from x_α and from all variables in A_α . Here D_β is $[\lambda y_\gamma B_\delta]$. Let y be an arbitrary member of \mathcal{D}_γ . Let $\varphi' = (\varphi : y_\gamma/y)$. Then $\mathcal{V}_\varphi A_\alpha = \mathcal{V}_{\varphi'} A_\alpha$ by 5400, and $(\varphi' : x_\alpha/\mathcal{V}_{\varphi'} A_\alpha) = (\psi : y_\gamma/y)$ since x_α and y_γ are distinct. Hence $(\mathcal{V}_\psi D_\beta)y = \mathcal{V}_{(\psi:y_\gamma/y)} B_\delta = \mathcal{V}_{\varphi'}[[\lambda x_\alpha B_\delta] A_\alpha] = (\mathcal{V}_\varphi[\lambda y_\gamma \bullet [\lambda x_\alpha B_\delta] A_\alpha])y$, so the functions $\mathcal{V}_\psi D_\beta$ and $\mathcal{V}_\varphi[\lambda y_\gamma \bullet [\lambda x_\alpha B_\delta] A_\alpha]$ are the same. Hence \models_φ [Axiom 4₄].

Axiom 4₅. $[\lambda x_\alpha \bullet \lambda x_\alpha B_\delta] A_\alpha = [\lambda x_\alpha B_\delta]$. Here D_β is $[\lambda x_\alpha B_\delta]$ and $\mathcal{V}_\psi D_\beta = \mathcal{V}_\varphi D_\beta$ by 5400 since x_α is not free in D_β , so \models_φ [Axiom 4₅].

Axiom 5. $\iota_{\iota(\alpha)}[Q_{\alpha}y] = y_i$. Clearly $\mathcal{V}_{\varphi}[\iota_{\iota(\alpha)} \bullet Q_{\alpha}y_i] = (\mathcal{V}\iota_{\iota(\alpha)})((\mathcal{V}Q_{\alpha})\mathcal{V}_{\varphi}y_i) = \mathcal{V}_{\varphi}y_i$ by the definitions of $\mathcal{V}\iota_{\iota(\alpha)}$ and $\mathcal{V}Q_{\alpha}$, so $\models_{\varphi} [\text{Axiom 5}]$.

It is now clear that every wff which occurs in a proof is valid in the general sense, so (a) is proved.

To prove (b), suppose $\mathcal{G} \vdash A$. Then there is a finite subset $\{H^1, \dots, H^n\}$ of \mathcal{G} such that $H^1, \dots, H^n \vdash A$, so by the Deduction Theorem $\vdash H^1 \supset \dots \supset H^n \supset A$, so by (a) $\mathcal{M} \models H^1 \supset \dots \supset H^n \supset A$. Also $\mathcal{M} \models H^i$ for $1 \leq i \leq n$ since \mathcal{M} is a model for \mathcal{G} , so $\mathcal{M} \models A$.

REMARK. We can now see that \mathcal{Q}_0 is a conservative extension of the system \mathcal{F} of first-order logic, in the sense that if C is any wff of \mathcal{F} and C' is the result of translating C into the notation of \mathcal{Q}_0 , then $\vdash_{\mathcal{F}} C$ if and only if $\vdash_{\mathcal{Q}_0} C'$. The proof in one direction follows directly from Rule Q (5236). For the other direction, suppose $\vdash_{\mathcal{Q}_0} C'$. Then $\models C'$ by 5402. It is not hard to see that every model of \mathcal{F} can be extended to a model of \mathcal{Q}_0 , and hence that C is valid in the sense of first-order logic. Therefore $\vdash_{\mathcal{F}} C$ by Gödel's Completeness Theorem (2510). Similar considerations show that \mathcal{F}^ω is a conservative extension of \mathcal{F} . In summary, every first-order theorem of type theory has a first-order proof.

However, some theorems of first-order logic can be proved most efficiently by using concepts which can be expressed only in higher-order logic. Examples may be found in [Andrews and Bishop, 1996] and [Boolos, 1998, Chapter 25]. Statman proved [Statman, 1978, Proposition 6.3.5] that the minimal length of a proof in first-order logic of a wff of first-order logic may be extraordinarily longer than that the minimal length of a proof of the same wff in second-order logic. A related result by Gödel [Gödel, 1936] is that in general “passing to the logic of the next higher order has the effect, not only of making provable certain propositions that were not provable before, but also of making it possible to shorten, by an extraordinary amount, infinitely many of the proofs already available”. A complete proof of this may be found in [Buss, 1994].

DEFINITION. A set \mathcal{G} of wffs_o is *inconsistent* iff $\mathcal{G} \vdash F_o$. \mathcal{Q}_0 is inconsistent iff $\vdash F_o$. As usual, “consistent” means “not inconsistent”. A wff_o B is *consistent with* \mathcal{G} iff $\mathcal{G} \cup \{B\}$ is consistent. Note that by VI (5215) and the definition of F_o , \mathcal{G} is inconsistent iff for every wff_o A , $\mathcal{G} \vdash A$.

5403 Consistency Theorem \mathcal{Q}_0 is consistent, and every set of wffs_o which has a model is consistent.

Proof: If \mathcal{G} is a set of wffs, which has a model \mathcal{M} but \mathcal{G} is inconsistent, then $\mathcal{G} \vdash F_o$ so $\mathcal{M} \models F_o$ by 5402, contradicting 5401(d). Hence a set of wffs which has a model must be consistent.

To show that \mathcal{Q}_0 is consistent, i.e., that not $\vdash F_o$, we must simply note that there is a model \mathcal{M} for the empty set of wffs, i.e., models exist. Let d be any object, and let \mathcal{M} be the standard model with $\mathcal{D}_i = \{d\}$. (Each of the domains \mathcal{D}_α is finite, and can be described very explicitly.) Note that $\mathcal{V}^{\mathcal{M}} \iota_{(o\alpha)}$ must be the constant function whose value is d for all arguments. ■

DEFINITION. A *countable model* is a model $\langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{J} \rangle$ in which each of the domains \mathcal{D}_α is countable (finite or denumerable). A *finite model* is a model in which each of the domains \mathcal{D}_α is finite.

Note that a model is finite iff \mathcal{D}_i is finite.

5404 Theorem. Every finite model for \mathcal{Q}_0 is standard.

*Proof:*⁵ Let $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{J} \rangle$ be any finite model for \mathcal{Q}_0 . For arbitrary type symbols α and β we must show that $\mathcal{D}_{\alpha\beta}$ contains all functions from \mathcal{D}_β to \mathcal{D}_α . So let g be any such function. Let $\mathcal{D}_\beta = \{m^1, \dots, m^k\}$. Let φ be an assignment with values on the variables $w_\beta^1, \dots, w_\beta^k, z_\alpha^1, \dots, z_\alpha^k$ as follows: $\varphi w_\beta^1 = m^1, \dots, \varphi w_\beta^k = m^k, \varphi z_\alpha^1 = gm^1, \dots, \varphi z_\alpha^k = gm^k$. Let A be the wff

$$[[x_\beta = w_\beta^1 \wedge y_\alpha = z_\alpha^1] \vee \dots \vee [x_\beta = w_\beta^k \wedge y_\alpha = z_\alpha^k]].$$

We shall show that $g = \mathcal{V}_\varphi [\lambda x_\beta \bullet \iota y_\alpha A]$, which is in $\mathcal{D}_{\alpha\beta}$ since \mathcal{M} is a general model.

We first prove a theorem of \mathcal{Q}_0 which will be useful. Let $1 \leq j \leq k$, and let P be the wff

$$[[x_\beta = w_\beta^j] \wedge \bigwedge_{i \in \{1, \dots, k\} - \{j\}} [x_\beta \neq w_\beta^i]]$$

- (.1) $.1 \vdash P$ hyp
- (.2) $.1 \vdash A = \bullet z_\alpha^j = y_\alpha$ Rule P: .1, 5302
- (.3) $.1 \vdash \iota y_\alpha A = \iota_{\alpha(o\alpha)} \lambda y_\alpha A$ 5200, def of ιy_α
- (.4) $.1 \vdash \iota y_\alpha A = \iota_{\alpha(o\alpha)} \lambda y_\alpha \bullet Q_{o\alpha\alpha} z_\alpha^j y_\alpha$ R' : .2, .3, def of $=$
(Note that y_α is not free in the hypothesis P .)
- (.5) $.1 \vdash \iota y_\alpha A = \iota_{\alpha(o\alpha)} \bullet Q_{o\alpha\alpha} z_\alpha^j$ R': 5205, .4

⁵The idea for this proof comes from [Henkin, 1963b].

$$\begin{aligned} (.6) \quad & .1 \vdash \exists y_\alpha A = z_\alpha^j \\ (.7) \quad & \vdash P \supset \exists y_\alpha A = z_\alpha^j \end{aligned}$$

R': 5309, .5
5240: .6

Now let m^j be an arbitrary element of \mathcal{D}_β , and let $\psi = (\varphi : x_\beta / m^j)$. Note that $\mathcal{M} \models_\psi P$, so $\mathcal{V}_\psi[\exists y_\alpha A] = \mathcal{V}_\psi z_\alpha^j$ by .7, 5402, and 5401. Therefore

$$(\mathcal{V}_\varphi[\lambda x_\beta \exists y_\alpha A])m^j = \mathcal{V}_\psi[\exists y_\alpha A] = \mathcal{V}_\psi z_\alpha^j = \varphi z_\alpha^j = gm^j.$$

Thus $\mathcal{V}_\varphi[\lambda x_\beta \exists y_\alpha A]$ and g have the same values for all arguments in \mathcal{D}_β , so they are the same function. ■

In 5504-5506 we shall see that many infinite models are nonstandard.

REMARK. It is easy to see that all instances of the Axiom Schema of Choice are true in every finite model, since the elements of each type can be enumerated, and we can define (in the meta-language) a function which chooses from each nonempty set the first member under this enumeration. Since (by 5404) the model is standard, this function will actually be in the model, so the Axiom of Choice will be true in the model. Note that the proof of the consistency of \mathcal{Q}_0 (5403) given above would still work if we added to \mathcal{Q}_0 the Axiom Schema of Choice. Of course, this result is rather trivial. Significant consistency questions arise only when an Axiom of Infinity is added to \mathcal{Q}_0 . Such an axiom will be introduced in Chapter 6.

DEFINITION. Let $\mathcal{N} = \langle \{\mathcal{E}_\alpha\}_\alpha, \mathcal{J} \rangle$ and $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{I} \rangle$ be general models of \mathcal{Q}_0 . A mapping τ is an *isomorphism* from \mathcal{N} to \mathcal{M} iff for all type symbols α and β ,

- (1) τ maps \mathcal{E}_α onto \mathcal{D}_α ;
- (2) $\tau T = T$ and $\tau F = F$;
- (3) For all $f \in \mathcal{E}_{\alpha\beta}$ and $x \in \mathcal{E}_\beta$, $(\tau f)(\tau x) = \tau(fx)$ (see Figure 5.1);
- (4) For each primitive constant c_α of \mathcal{Q}_0 , $\tau \mathcal{J} c_\alpha = \mathcal{I} c_\alpha$.

\mathcal{N} and \mathcal{M} are *isomorphic* iff there is an isomorphism from \mathcal{N} to \mathcal{M} .

5405 Proposition. An isomorphism of general models is a one-to-one mapping.

Proof: We use the notation in the definition above. Suppose τ is an isomorphism, and x and y are distinct elements of \mathcal{E}_α . We shall show that $\tau x \neq \tau y$.

$$\begin{array}{ccccc}
 x \in \mathcal{E}_\beta & \xrightarrow{\tau} & \mathcal{D}_\beta & \tau x \\
 f \in \mathcal{E}_{\alpha\beta} & \downarrow & & \downarrow \tau f \in \mathcal{D}_{\alpha\beta} \\
 fx \in \mathcal{E}_\alpha & \xrightarrow[\tau]{} & \mathcal{D}_\alpha & (\tau f)(\tau x) = \tau(fx)
 \end{array}$$

Figure 5.1: Isomorphism condition (3)

Let q be the identity relation on \mathcal{E}_α . $\tau q = \tau(\mathcal{J}\mathbf{Q}_{o\alpha\alpha}) = \mathcal{I}\mathbf{Q}_{o\alpha\alpha}$, which is the identity relation on \mathcal{D}_α , and

$$((\tau q)(\tau x))(\tau y) = (\tau(qx))(\tau y) \quad (\text{by 3})$$

$$= \tau(qxy) \quad (\text{by 3})$$

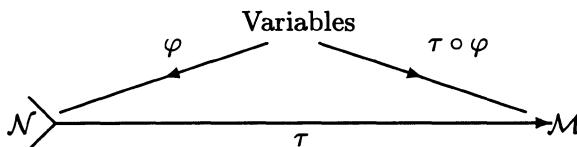
$$= \tau F \quad (\text{since } x \neq y)$$

$$= F \quad (\text{by 2})$$

so $\tau x \neq \tau y$. ■

It is readily verified that the relation of being isomorphic is an equivalence relation (X5404).

5406 Proposition. Let \mathcal{N} and \mathcal{M} be general models for \mathcal{Q}_0 , and let τ be an isomorphism from \mathcal{N} to \mathcal{M} . For each assignment φ into \mathcal{N} , let $\tau \circ \varphi$ be the assignment into \mathcal{M} such that for each variable x_γ , $(\tau \circ \varphi)x_\gamma = \tau(\varphi x_\gamma)$. Then for every wff C_γ of \mathcal{Q}_0 , $\tau(\mathcal{V}_\varphi^\mathcal{N} C_\gamma) = \mathcal{V}_{\tau \circ \varphi}^\mathcal{M} C_\gamma$. (See Figure 5.2.)

Figure 5.2: Assignments into \mathcal{N} and \mathcal{M}

Proof: Suppose $\mathcal{N} = \langle \{\mathcal{E}_\alpha\}_\alpha, \mathcal{J} \rangle$ and $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{I} \rangle$. We prove the proposition for all assignments φ by induction on the construction of C_γ .

- (a) If C_γ is a variable x_γ , then $\tau(\mathcal{V}_\varphi^\mathcal{N} C_\gamma) = \tau(\varphi x_\gamma) = (\tau \circ \varphi)x_\gamma = \mathcal{V}_{\tau \circ \varphi}^\mathcal{M} C_\gamma$.
- (b) If C_γ is a primitive constant of \mathcal{Q}_0 , then $\tau(\mathcal{V}_\varphi^\mathcal{N} C_\gamma) = \tau(\mathcal{J}C_\gamma) = \mathcal{I}C_\gamma = \mathcal{V}_{\tau \circ \varphi}^\mathcal{M} C_\gamma$.

- (c) If $C_\gamma = [A_{\gamma\beta}B_\beta]$, then $\tau(\mathcal{V}_\varphi^N C_\gamma) = \tau((\mathcal{V}_\varphi^N A_{\gamma\beta})(\mathcal{V}_\varphi^N B_\beta))$
 $= (\tau\mathcal{V}_\varphi^N A_{\gamma\beta}) (\tau\mathcal{V}_\varphi^N B_\beta)$
 $= (\mathcal{V}_{\tau\circ\varphi}^M A_{\gamma\beta})(\mathcal{V}_{\tau\circ\varphi}^M B_\beta)$ (by inductive hypothesis)
 $= \mathcal{V}_{\tau\circ\varphi}^M [A_{\gamma\beta}B_\beta].$
- (d) Suppose $C_\gamma = [\lambda x_\beta A_\alpha]$. Let z be an arbitrary member of D_β . Let $w = \tau^{-1}z$, so $w \in E_\beta$ and $\tau w = z$. It is readily checked that $\tau \circ (\varphi : x_\beta/w) = (\tau \circ \varphi : x_\beta/z)$.
 $(\tau\mathcal{V}_\varphi^N [\lambda x_\beta A_\alpha])z = \tau((\mathcal{V}_\varphi^N [\lambda x_\beta A_\alpha])w) = \tau(\mathcal{V}_{(\varphi:x_\beta/w)}^N A_\alpha)$
 $= \mathcal{V}_{\tau\circ(\varphi:x_\beta/w)}^M A_\alpha$ (by inductive hypothesis)
 $= \mathcal{V}_{(\tau\circ\varphi:x_\beta/z)}^M A_\alpha = (\mathcal{V}_{\tau\circ\varphi}^M [\lambda x_\beta A_\alpha])z.$

Since this is true for each $z \in D_\beta$, $\tau\mathcal{V}_\varphi^N [\lambda x_\beta A_\alpha] = \mathcal{V}_{\tau\circ\varphi}^M [\lambda x_\beta A_\alpha]$. ■

5407 Corollary. If N and M are isomorphic general models of \mathcal{Q}_0 , then $M \models C$ iff $N \models C$.

Proof: Suppose $M \models C$, and τ is an isomorphism from N to M . Then for any assignment φ into N , since τ is the identity function on D_o , $\mathcal{V}_\varphi^N C = \tau(\mathcal{V}_\varphi^N C) = \mathcal{V}_{\tau\circ\varphi}^M C = T$, so $N \models C$.

Similarly, if $N \models C$, then $M \models C$. ■

EXERCISES

X5400. Prove that $\models 5207$ without using 5402.

X5401. Show that not every wff of the form $[A_\alpha = B_\alpha] \supset \neg [\lambda z_\beta A_\alpha] = [\lambda z_\beta B_\alpha]$ is valid.

X5402. Show that every wff of the form $[A_\alpha = B_\alpha] \supset \neg [\lambda z_\beta A_\alpha] = [\lambda z_\beta B_\alpha]$ is satisfiable in every general model.

X5403. Show that if M is any general model, φ is any assignment into M , and A is any wff_o, then $M \models_\varphi \exists x_\alpha A$ iff there is an assignment ψ which agrees with φ off x_α such that $M \models_\psi A$.

X5404. Show that the relation of being isomorphic is an equivalence relation between models of \mathcal{Q}_0 .

X5405. Prove that any theorem of \mathcal{Q}_0 in β -normal form has the form $[A_\alpha = B_\alpha]$. (Therefore, by Exercise X5213, every theorem of \mathcal{Q}_0 can be transformed to the form $[A_\alpha = B_\alpha]$ by a sequence of β -contractions and α -conversions.)

X5406. Let \mathcal{B} be the system which is obtained from \mathcal{Q}_0 when Axiom Schema 2^α is replaced by the schema $[x_\alpha = y_\alpha] = \forall h_{o\alpha} [h_{o\alpha} x_\alpha = h_{o\alpha} y_\alpha]$. Is 2^α a theorem schema of \mathcal{B} ?

§55. Completeness of \mathcal{Q}_0

Formulations of \mathcal{Q}_0 differ from one another in having different primitive constants. We denote by $\mathcal{L}(\mathcal{Q}_0)$ the *language* of \mathcal{Q}_0 , i.e., the set of wffs of \mathcal{Q}_0 . If \mathcal{Q}_0^1 and \mathcal{Q}_0^2 are formulations of \mathcal{Q}_0 such that every primitive symbol of \mathcal{Q}_0^1 is a primitive symbol of \mathcal{Q}_0^2 , we say that \mathcal{Q}_0^2 is an *expansion* of \mathcal{Q}_0^1 . Clearly, \mathcal{Q}_0^2 is an expansion of \mathcal{Q}_0^1 iff $\mathcal{L}(\mathcal{Q}_0^1) \subseteq \mathcal{L}(\mathcal{Q}_0^2)$.

We denote the cardinality of any set \mathcal{S} by $\text{card}(\mathcal{S})$. Since every formulation of \mathcal{Q}_0 contains the denumerable set of logical constants, and every wff is a finite sequence of primitive symbols, it is easy to see that $\text{card}(\mathcal{L}(\mathcal{Q}_0))$ is the cardinality of the set of primitive constants of \mathcal{Q}_0 . For most purposes we shall deal with formulations of \mathcal{Q}_0 in which the set of nonlogical primitive constants is countable (denumerably infinite or finite, and possibly empty), so that $\text{card}(\mathcal{L}(\mathcal{Q}_0)) = \aleph_0$, and we shall call such formulations of \mathcal{Q}_0 countable. Indeed, unless otherwise indicated, the reader may usually assume that we are dealing with a formulation of \mathcal{Q}_0 which has no nonlogical constants. However, we shall impose no actual restrictions on the cardinality of the set of nonlogical constants.

DEFINITIONS: Let \mathcal{H} be a set of sentences of \mathcal{Q}_0 .

- (1) \mathcal{H} is *complete* in \mathcal{Q}_0 iff for every sentence \mathbf{A} of \mathcal{Q}_0 , either $\mathcal{H} \vdash \mathbf{A}$ or $\mathcal{H} \vdash \sim \mathbf{A}$.
- (2) \mathcal{H} is *extensionally complete* in \mathcal{Q}_0 iff for every sentence of \mathcal{Q}_0 of the form $[\mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}]$, there is a cwff \mathbf{C}_β of \mathcal{Q}_0 such that

$$\mathcal{H} \vdash [\mathbf{A}_{\alpha\beta} \mathbf{C}_\beta = \mathbf{B}_{\alpha\beta} \mathbf{C}_\beta] \supset \neg \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}. \quad (*)$$

REMARK. The wff $(*)$ is equivalent to $\mathbf{A}_{\alpha\beta} \neq \mathbf{B}_{\alpha\beta} \supset \neg \mathbf{A}_{\alpha\beta} \mathbf{C}_\beta \neq \mathbf{B}_{\alpha\beta} \mathbf{C}_\beta$, which says that if $\mathbf{A}_{\alpha\beta}$ and $\mathbf{B}_{\alpha\beta}$ are different functions, then \mathbf{C}_β is an argument on which they differ. Note that a special case of the wff schema $(*)$ is $[\lambda x_\beta T] \mathbf{C}_\beta = [\lambda x_\beta \mathbf{D}] \mathbf{C}_\beta \supset \neg [\lambda x_\beta T] = [\lambda x_\beta \mathbf{D}]$, which is equivalent to $[\lambda x_\beta \mathbf{D}] \mathbf{C}_\beta \supset \forall x_\beta \mathbf{D}$. If \mathbf{D} is $\sim \mathbf{E}$. This is equivalent to $\exists x_\beta \mathbf{E} \supset [\lambda x_\beta \mathbf{E}] \mathbf{C}_\beta$.

REMARK. If from \mathcal{H} one can derive the Axiom of Choice (for each type symbol β) in a form such as $\forall p_{o\beta} \exists x_\beta p_{o\beta} x_\beta \supset p_{o\beta} \neg \mathbf{C}_{\beta(o\beta)} p_{o\beta}$ (for some

cwff $\mathbf{C}_{\beta(o\beta)}$), then \mathcal{H} is extensionally complete, since

$$\begin{aligned}\mathcal{H} \vdash & \forall f_{\alpha\beta} \forall g_{\alpha\beta} \bullet f_{\alpha\beta}[\mathbf{C}_{\beta(o\beta)} \bullet \lambda x_\beta \bullet f_{\alpha\beta}x_\beta \neq g_{\alpha\beta}x_\beta] \\ & = g_{\alpha\beta}[\mathbf{C}_{\beta(o\beta)} \bullet \lambda x_\beta \bullet f_{\alpha\beta}x_\beta \neq g_{\alpha\beta}x_\beta] \supset f_{\alpha\beta} = g_{\alpha\beta}.\end{aligned}$$

(See Exercise X5500 below.)

REMARK. Some examples of complete sets of sentences are $\{\mathbf{A} \mid \mathcal{M} \models \mathbf{A}$ and \mathbf{A} is a sentence} for any particular model \mathcal{M} , $\{\forall x_i \forall y_i \bullet x_i = y_i\}$ (see Exercise X8032), and $\{F_o\}$.

5500 Extension Lemma. Let \mathcal{G} be any consistent set of sentences of \mathcal{Q}_0 . Then there is an expansion \mathcal{Q}_0^+ of \mathcal{Q}_0 and a set of \mathcal{H} of sentences of \mathcal{Q}_0^+ such that

- (1) $\mathcal{G} \subseteq \mathcal{H}$.
- (2) \mathcal{H} is consistent.
- (3) \mathcal{H} is complete in \mathcal{Q}_0^+ .
- (4) \mathcal{H} is extensionally complete in \mathcal{Q}_0^+ .
- (5) $\text{card}(\mathcal{L}(\mathcal{Q}_0^+)) = \text{card}(\mathcal{L}(\mathcal{Q}))$.

Proof: Let $K = \text{card}(\mathcal{L}(\mathcal{Q}_0))$ and for each type symbol α let \mathcal{C}_α be a well-ordered set of cardinality K of new primitive constants of type α , and let $\mathcal{C} = \bigcup_\alpha \mathcal{C}_\alpha$ (the union being taken over all type symbols α). Let \mathcal{Q}_0^+ be the expansion of \mathcal{Q}_0 obtained by adding the constants in \mathcal{C} to the primitive symbols of \mathcal{Q}_0 . Clearly $\text{card}(\mathcal{L}(\mathcal{Q}_0^+)) = K$.

Let κ be the initial ordinal of cardinality K . Well-order the sentences of \mathcal{Q}_0^+ , and for each ordinal $\tau < \kappa$, let \mathbf{S}^τ be the τ th sentence of \mathcal{Q}_0^+ under this well-ordering.⁶

By transfinite induction, for each ordinal $\tau \leq \kappa$ we define a set \mathcal{G}_τ of sentences of \mathcal{Q}_0^+ , so that $\mathcal{G}_\sigma \subseteq \mathcal{G}_\tau$ whenever $\sigma \leq \tau$, and there is a finite cardinal n such that at most $n + \text{card}(\tau)$ constants from \mathcal{C} occur in the sentences of \mathcal{G}_τ . $\mathcal{G}_0 = \mathcal{G}$. If τ is a limit ordinal, $\mathcal{G}_\tau = \bigcup_{\sigma < \tau} \mathcal{G}_\sigma$. The definition for successor ordinals $\tau + 1$ is by cases:

⁶Very explicit well-orderings of sentences can easily be specified for countable formulations of \mathcal{Q}_0 . The method of assigning numbers to wffs given in §70 can easily be modified to apply to any formulation of \mathcal{Q}_0 with countably many parameters. The natural ordering of the numbers assigned to the sentences induces an ordering (which is a well-ordering) of the sentences.

- (a) If \mathbf{S}^τ is consistent with \mathcal{G}_τ , then $\mathcal{G}_{\tau+1} = \mathcal{G}_\tau \cup \{\mathbf{S}^\tau\}$.
- (b) If \mathbf{S}^τ is not consistent with \mathcal{G}_τ , and \mathbf{S}^τ is not of the form $[\mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}]$, then $\mathcal{G}_{\tau+1} = \mathcal{G}_\tau$.
- (c) If \mathbf{S}^τ is not consistent with \mathcal{G}_τ and \mathbf{S}^τ is of the form $[\mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}]$, then $\mathcal{G}_{\tau+1} = \mathcal{G}_\tau \cup \{\sim \cdot \mathbf{A}_{\alpha\beta} \mathbf{c}_\beta = \mathbf{B}_{\alpha\beta} \mathbf{c}_\beta\}$, where \mathbf{c}_β is the first constant in \mathcal{C}_β which does not occur in \mathcal{G}_τ or \mathbf{S}^τ . (Note that there must be such a constant.)

We next prove by induction on τ that each \mathcal{G}_τ is consistent in \mathcal{Q}_0^+ .

To prove that \mathcal{G}_0 is consistent in \mathcal{Q}_0^+ , suppose \mathcal{P} is a proof of F_o in \mathcal{Q}_0^+ from some finite subset \mathcal{S} of \mathcal{G}_0 . Let θ be a substitution which replaces the constants from \mathcal{C} which occur in \mathcal{P} by distinct variables of the same type which do not occur in \mathcal{P} . It is easily seen that $\theta\mathcal{P}$ is a proof in \mathcal{Q}_0 of F_o from \mathcal{S} , since no constant of \mathcal{C} occurs in \mathcal{G}_0 . This contradicts the consistency of \mathcal{G}_0 in \mathcal{Q}_0 .

Next, we suppose \mathcal{G}_τ is consistent and prove $\mathcal{G}_{\tau+1}$ consistent. In case (a) or (b) this is trivial. Suppose $\mathcal{G}_{\tau+1}$ is obtained by case (c) and is inconsistent. Thus $\mathcal{G}_{\tau+1} = \mathcal{G}_\tau \cup \{\sim \cdot \mathbf{A}_{\alpha\beta} \mathbf{c}_\beta = \mathbf{B}_{\alpha\beta} \mathbf{c}_\beta\}$ so by the Deduction Theorem and Rule P, $\mathcal{G}_\tau \vdash \mathbf{A}_{\alpha\beta} \mathbf{c}_\beta = \mathbf{B}_{\alpha\beta} \mathbf{c}_\beta$. Let \mathcal{P} be a proof of $\mathbf{A}_{\alpha\beta} \mathbf{c}_\beta = \mathbf{B}_{\alpha\beta} \mathbf{c}_\beta$ from a finite subset \mathcal{S} of \mathcal{G}_τ , and let \mathbf{x}_β be a variable of type β which does not occur in \mathcal{P} or \mathcal{S} . Since \mathbf{c}_β does not occur in \mathcal{G}_τ , $\mathbf{A}_{\alpha\beta}$, or $\mathbf{B}_{\alpha\beta}$, it is easily seen that $\mathcal{S}^{\mathbf{c}_\beta} \mathcal{P}$ (the result of substituting \mathbf{x}_β for \mathbf{c}_β throughout \mathcal{P}) is a proof from \mathcal{S} of $\mathbf{A}_{\alpha\beta} \mathbf{x}_\beta = \mathbf{B}_{\alpha\beta} \mathbf{x}_\beta$. Therefore

$$\begin{aligned} \mathcal{S} \vdash \mathbf{A}_{\alpha\beta} \mathbf{x}_\beta &= \mathbf{B}_{\alpha\beta} \mathbf{x}_\beta \\ \mathcal{S} \vdash \forall \mathbf{x}_\beta \cdot \mathbf{A}_{\alpha\beta} \mathbf{x}_\beta &= \mathbf{B}_{\alpha\beta} \mathbf{x}_\beta && \text{by Gen, since } \mathbf{x}_\beta \text{ does not occur in } \mathcal{S} \\ \mathcal{S} \vdash \mathbf{A}_{\alpha\beta} &= \mathbf{B}_{\alpha\beta} && \text{by RR and Axiom 3} \end{aligned}$$

Thus $\mathcal{G}_\tau \vdash \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$. But by the definition of $\mathcal{G}_{\tau+1}$ in case (c), the wff $[\mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}]$ is inconsistent with \mathcal{G}_τ , so $\mathcal{G}_\tau, [\mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}] \vdash F$, so \mathcal{G}_τ is inconsistent. This contradiction shows that $\mathcal{G}_{\tau+1}$ must be consistent when \mathcal{G}_τ is.

Finally, if τ is a limit ordinal and \mathcal{G}_σ is consistent for each $\sigma < \tau$, then \mathcal{G}_τ must be consistent, since each finite subset of \mathcal{G}_τ is a subset of some \mathcal{G}_σ with $\sigma < \tau$.

Let $\mathcal{H} = \mathcal{G}_\kappa$. Clearly \mathcal{H} is a consistent set of sentences of \mathcal{Q}_0^+ , and $\mathcal{G} \subseteq \mathcal{H}$.

To see that \mathcal{H} is complete and extensionally complete, let \mathbf{S} be any sentence of \mathcal{Q}_0^+ . Then there exists $\tau < \kappa$ such that \mathbf{S} is \mathbf{S}^τ . If \mathbf{S} is consistent with \mathcal{G}_τ then by case (a) $\mathbf{S} \in \mathcal{G}_{\tau+1} \subseteq \mathcal{H}$ so $\mathcal{H} \vdash \mathbf{S}$. Otherwise $\mathcal{G}_\tau, \mathbf{S} \vdash F$ so $\mathcal{G}_\tau \vdash \sim \mathbf{S}$ and $\mathcal{H} \vdash \sim \mathbf{S}$. Thus \mathcal{H} is complete. Moreover, suppose \mathbf{S} has the

form $\mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$; in the case where \mathbf{S} is consistent with \mathcal{G}_r , by Rule P $\mathcal{H} \vdash [\mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{C}_\beta] \supset \mathbf{S}$ for every wff \mathbf{C}_β ; in the case where \mathbf{S} is inconsistent with \mathcal{G}_r , by (c) there is a $\mathbf{c}_\beta \in \mathcal{C}$ such that $\mathcal{H} \vdash \sim \mathbf{A}_{\alpha\beta}\mathbf{c}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{c}_\beta$, so by Rule P, $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta}\mathbf{c}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{c}_\beta \supset \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$. ■

We shall say that a model $\langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{J} \rangle$ for \mathcal{Q}_0 is *frugal* iff for each type symbol α , $\text{card}(\mathcal{D}_\alpha) \leq \text{card}(\mathcal{L}(\mathcal{Q}_0))$. Thus, in the usual case where \mathcal{Q}_0 is countable, a *frugal* model is a *countable* model. Note that by the same reasoning as was used in §25 when the concept of frugality was introduced, if one has an uncountable formulation of \mathcal{Q}_0 , one cannot expect that every consistent set of sentences will have a countable model.

5501 Henkin's Theorem. Every consistent set of sentences of \mathcal{Q}_0 has a frugal general model.

Proof: Let \mathcal{G} be a consistent set of sentences of \mathcal{Q}_0 , and let \mathcal{H} and \mathcal{Q}_0^+ be as described in Lemma 5500.

REMARK. If \mathbf{A}_α and \mathbf{B}_α are cwffs of \mathcal{Q}_0^+ , we shall say that \mathbf{A}_α is *equivalent* to \mathbf{B}_α iff $\mathcal{H} \vdash \mathbf{A}_\alpha = \mathbf{B}_\alpha$. Clearly this is an equivalence relation. In essence, we shall take the equivalence classes of this relation to be the elements in our general model. However, an equivalence class of cwffs is not really a function, so we define a map \mathcal{V} such that $\mathcal{V}\mathbf{A}_{\alpha\beta}$ is a function which corresponds, in an appropriate sense, to the equivalence class containing $\mathbf{A}_{\alpha\beta}$.

We define, by induction on γ , the domains \mathcal{D}_γ and a function \mathcal{V} whose domain is the set of cwffs of \mathcal{Q}_0^+ , so that

- (1 γ) $\mathcal{D}_\gamma = \{\mathcal{V}\mathbf{A}_\gamma \mid \mathbf{A}_\gamma \text{ is a cwff}_\gamma\}$ and
- (2 γ) for all cwffs \mathbf{A}_γ and \mathbf{B}_γ , $\mathcal{V}\mathbf{A}_\gamma = \mathcal{V}\mathbf{B}_\gamma$ iff $\mathcal{H} \vdash \mathbf{A}_\gamma = \mathbf{B}_\gamma$.

Let $\mathcal{D}_o = \{\mathsf{T}, \mathsf{F}\}$, and for each sentence \mathbf{A} , let $\mathcal{V}\mathbf{A} = \mathsf{T}$ if $\mathcal{H} \vdash \mathbf{A}$; otherwise let $\mathcal{V}\mathbf{A} = \mathsf{F}$. Since \mathcal{H} is complete and consistent, exactly one of the conditions $\mathcal{H} \vdash \mathbf{A}$ and $\mathcal{H} \vdash \sim \mathbf{A}$ holds for each sentence \mathbf{A} , so (1 \circ) and (2 \circ) are satisfied.

For each cwff, \mathbf{A}_i , let $\mathcal{V}\mathbf{A}_i = \{\mathbf{B}_i \mid \mathbf{B}_i \text{ is a cwff}_i \text{ and } \mathcal{H} \vdash \mathbf{A}_i = \mathbf{B}_i\}$, and let $\mathcal{D}_i = \{\mathcal{V}\mathbf{A}_i \mid \mathbf{A}_i \text{ is a cwff}_i\}$. Clearly $\mathcal{V}\mathbf{A}_i = \mathcal{V}\mathbf{B}_i$ iff $\mathcal{H} \vdash \mathbf{A}_i = \mathbf{B}_i$.

To define $\mathcal{D}_{\alpha\beta}$, suppose \mathcal{D}_α and \mathcal{D}_β are defined, and $\mathbf{A}_{\alpha\beta}$ is a cwff $_{\alpha\beta}$. Let $\mathcal{V}\mathbf{A}_{\alpha\beta}$ be the function from \mathcal{D}_β to \mathcal{D}_α whose value, for any argument $\mathbf{VC}_\beta \in \mathcal{D}_\beta$, is $\mathcal{V}[\mathbf{A}_{\alpha\beta}\mathbf{C}_\beta]$. To show that this definition is unambiguous, we must show that it is independent of the particular cwff \mathbf{C}_β used to represent the argument. So suppose $\mathcal{V}\mathbf{B}_\beta = \mathcal{V}\mathbf{C}_\beta$; then $\mathcal{H} \vdash \mathbf{B}_\beta = \mathbf{C}_\beta$, so $\mathcal{H} \vdash$

$\mathbf{A}_{\alpha\beta}\mathbf{B}_\beta = \mathbf{A}_{\alpha\beta}\mathbf{C}_\beta$, so $\mathcal{V}[\mathbf{A}_{\alpha\beta}\mathbf{B}_\beta] = \mathcal{V}[\mathbf{A}_{\alpha\beta}\mathbf{C}_\beta]$. Finally, let $\mathcal{D}_{\alpha\beta} = \{\mathcal{V}\mathbf{A}_{\alpha\beta} \mid \mathbf{A}_{\alpha\beta} \text{ is a cwff}_{\alpha\beta}\}$.

We must show that $(2^{\alpha\beta})$ is satisfied. If $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$, then for all cwffs \mathbf{C}_β , $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{C}_\beta$, so for all arguments $\mathcal{V}\mathbf{C}_\beta$, $(\mathcal{V}\mathbf{A}_{\alpha\beta})(\mathcal{V}\mathbf{C}_\beta) = \mathcal{V}[\mathbf{A}_{\alpha\beta}\mathbf{C}_\beta] = \mathcal{V}[\mathbf{B}_{\alpha\beta}\mathbf{C}_\beta] = (\mathcal{V}\mathbf{B}_{\alpha\beta})(\mathcal{V}\mathbf{C}_\beta)$, so $\mathcal{V}\mathbf{A}_{\alpha\beta} = \mathcal{V}\mathbf{B}_{\alpha\beta}$. On the other hand, suppose $\mathcal{V}\mathbf{A}_{\alpha\beta} = \mathcal{V}\mathbf{B}_{\alpha\beta}$, and let \mathbf{C}_β be a cwff such that $\mathcal{H} \vdash [\mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{C}_\beta \supset \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}]$; then $\mathcal{V}[\mathbf{A}_{\alpha\beta}\mathbf{C}_\beta] = (\mathcal{V}\mathbf{A}_{\alpha\beta})(\mathcal{V}\mathbf{C}_\beta) = (\mathcal{V}\mathbf{B}_{\alpha\beta})(\mathcal{V}\mathbf{C}_\beta) = \mathcal{V}[\mathbf{B}_{\alpha\beta}\mathbf{C}_\beta]$, so $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{C}_\beta$, so $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$.

Next we shall show that $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{V} \rangle$ constitutes an interpretation. Clearly \mathcal{V} maps each primitive constant of \mathcal{Q}_0^+ of type γ into \mathcal{D}_γ . We must show that $\mathcal{V}Q_{o\alpha\alpha}$ is the identity relation on \mathcal{D}_α . Let $\mathcal{V}\mathbf{A}_\alpha$ and $\mathcal{V}\mathbf{B}_\alpha$ be arbitrary members of \mathcal{D}_α . Then $\mathcal{V}\mathbf{A}_\alpha = \mathcal{V}\mathbf{B}_\alpha$ iff $\mathcal{H} \vdash \mathbf{A}_\alpha = \mathbf{B}_\alpha$ iff $\mathcal{H} \vdash Q_{o\alpha\alpha}\mathbf{A}_\alpha\mathbf{B}_\alpha$ iff $\mathsf{T} = \mathcal{V}[Q_{o\alpha\alpha}\mathbf{A}_\alpha\mathbf{B}_\alpha] = (\mathcal{V}Q_{o\alpha\alpha})(\mathcal{V}\mathbf{A}_\alpha)(\mathcal{V}\mathbf{B}_\alpha)$, so $\mathcal{V}Q_{o\alpha\alpha}$ is indeed the identity relation on \mathcal{D}_α .

We must show that $\mathcal{V}\iota_{\iota(o)}$ maps one-element sets in \mathcal{D}_o to their unique members. Let $\mathcal{V}\mathbf{A}_i$ be an arbitrary member of \mathcal{D}_i . Since $\mathcal{V}Q_{on}$ is the identity relation on \mathcal{D}_n , $((\mathcal{V}Q_{on})\mathcal{V}\mathbf{A}_i)$ is the one-element set whose sole member is $(\mathcal{V}\mathbf{A}_i)$. $\mathcal{H} \vdash \mathbf{A}_i = \iota_{\iota(o)}[Q_{on}\mathbf{A}_i]$ by Axiom 5, so $\mathcal{V}\mathbf{A}_i = \mathcal{V}(\iota_{\iota(o)}[Q_{on}\mathbf{A}_i]) = (\mathcal{V}\iota_{\iota(o)})((\mathcal{V}Q_{on})\mathcal{V}\mathbf{A}_i)$, as required.

Thus \mathcal{M} is indeed an interpretation, and we proceed to show that it is a general model for \mathcal{Q}_0^+ . For each assignment φ into \mathcal{M} and wff \mathbf{C}_γ , let $\mathbf{C}_\gamma^\varphi = S_{\mathbf{E}_{\delta_1}^1 \dots \mathbf{E}_{\delta_n}^n}^{x_{\delta_1}^1 \dots x_{\delta_n}^n} \mathbf{C}_\gamma$, where $x_{\delta_1}^1 \dots x_{\delta_n}^n$ are the free variables of \mathbf{C}_γ , and (for $1 \leq i \leq n$) $\mathbf{E}_{\delta_i}^i$ is the first cwff (in some fixed enumeration) of \mathcal{Q}_0^+ such that $\varphi x_{\delta_i}^i = \mathcal{V}\mathbf{E}_{\delta_i}^i$. Let $\mathcal{V}_\varphi \mathbf{C}_\gamma = \mathcal{V}\mathbf{C}_\gamma^\varphi$. Clearly $\mathbf{C}_\gamma^\varphi$ is a cwff, so $\mathcal{V}_\varphi \mathbf{C}_\gamma \in \mathcal{D}_\gamma$.

- (a) For any variable \mathbf{x}_δ , choose \mathbf{E}_δ so that $\varphi \mathbf{x}_\delta = \mathcal{V}\mathbf{E}_\delta$ as above; then $\mathcal{V}_\varphi \mathbf{x}_\delta = \mathcal{V}\mathbf{x}_\delta^\varphi = \mathcal{V}\mathbf{E}_\delta = \varphi \mathbf{x}_\delta$.
- (b) For each cwff (and hence for each primitive constant) \mathbf{C}_γ , $\mathcal{V}_\varphi \mathbf{C}_\gamma = \mathcal{V}\mathbf{C}_\gamma^\varphi = \mathcal{V}\mathbf{C}_\gamma$.
- (c) $\mathcal{V}_\varphi [\mathbf{A}_{\alpha\beta}\mathbf{B}_\beta] = \mathcal{V}[\mathbf{A}_{\alpha\beta}^\varphi\mathbf{B}_\beta^\varphi] = (\mathcal{V}\mathbf{A}_{\alpha\beta}^\varphi)(\mathcal{V}\mathbf{B}_\beta^\varphi) = (\mathcal{V}_\varphi \mathbf{A}_{\alpha\beta})(\mathcal{V}_\varphi \mathbf{B}_\beta)$.
- (d) Let $\mathcal{V}\mathbf{E}_\delta$ be an arbitrary member of \mathcal{D}_δ ; we may assume \mathbf{E}_δ is the first cwff which represents this member of \mathcal{D}_δ . Given an assignment φ and a variable \mathbf{x}_δ , let $\psi = (\varphi : \mathbf{x}_\delta / \mathcal{V}\mathbf{E}_\delta)$. It can be seen with the aid of 5207 that $\vdash [\lambda \mathbf{x}_\delta \mathbf{B}_\beta]^\varphi \mathbf{E}_\delta = \mathbf{B}_\beta^\psi$, so $(\mathcal{V}_\varphi [\lambda \mathbf{x}_\delta \mathbf{B}_\beta])(\mathcal{V}\mathbf{E}_\delta) = (\mathcal{V}[\lambda \mathbf{x}_\delta \mathbf{B}_\beta]^\varphi)(\mathcal{V}\mathbf{E}_\delta) = \mathcal{V}([\lambda \mathbf{x}_\delta \mathbf{B}_\beta]^\varphi \mathbf{E}_\delta) = \mathcal{V}\mathbf{B}_\beta^\psi = \mathcal{V}_\psi \mathbf{B}_\beta$. Thus $\mathcal{V}_\varphi [\lambda \mathbf{x}_\delta \mathbf{B}_\beta]$ satisfies condition (d) in the definition of a general model.

Thus \mathcal{M} is a general model for \mathcal{Q}_0^+ (and hence for \mathcal{Q}_0). Also, if $\mathbf{A} \in \mathcal{G}$, then $\mathbf{A} \in \mathcal{H}$, so $\mathcal{H} \vdash \mathbf{A}$, so $\mathcal{V}\mathbf{A} = \mathsf{T}$ and $\mathcal{M} \models \mathbf{A}$, so \mathcal{M} is a general model

for \mathcal{G} . Clearly each \mathcal{D}_α has cardinality $\leq \text{card}(\mathcal{L}(\mathcal{Q}_0))$, since \mathcal{V} maps the set of cwffs $_\alpha$ of \mathcal{Q}_0^+ onto \mathcal{D}_α , and $\text{card}(\mathcal{L}(\mathcal{Q}_0^+)) = \text{card}(\mathcal{L}(\mathcal{Q}_0))$. ■

5502 Henkin's Completeness and Soundness Theorem. Let \mathbf{A} be a wff $_o$ and let \mathcal{G} be a set of sentences.

- (a) $\vdash \mathbf{A}$ iff $\models \mathbf{A}$.
- (b) $\mathcal{G} \vdash \mathbf{A}$ iff \mathbf{A} is valid in every general model for \mathcal{G} .
- (c) $\mathcal{G} \vdash \mathbf{A}$ iff \mathbf{A} is valid in every frugal general model for \mathcal{G} .

Proof: (a) is a special case of (b) with $\mathcal{G} = \emptyset$, and (b) follows from (c) and the Soundness Theorem (5402), so we prove only (c).

Suppose \mathbf{A} is valid in every frugal general model for \mathcal{G} . Let \mathbf{B} be a universal closure of \mathbf{A} . Then \mathbf{B} is valid in every frugal general model of \mathcal{G} by 5401(g). Suppose $\mathcal{G} \cup \{\sim \mathbf{B}\}$ is consistent. Then by Henkin's Theorem, this set of sentences has a frugal general model \mathcal{M} , and $\mathcal{M} \models \sim \mathbf{B}$. But \mathcal{M} is a frugal model for \mathcal{G} , so $\mathcal{M} \models \mathbf{B}$. This contradiction shows that $\mathcal{G} \cup \{\sim \mathbf{B}\}$ must be inconsistent, so $\mathcal{G} \vdash \mathbf{B}$ by the Deduction Theorem and Rule P. Hence $\mathcal{G} \vdash \mathbf{A}$ by VI(5215). This, with the Soundness Theorem, completes the Proof. ■

The reader may have noticed a similarity between the basic approaches used to prove 5502 (Henkin's Completeness Theorem for \mathcal{Q}_0) and 2509 (the Generalized Completeness Theorem for first order logic). However, the approach to 2509 was carried out in a more abstract setting, since we have not presented anything analogous to Smullyan's Unifying Principle for \mathcal{Q}_0 . Actually, an extension of Smullyan's Unifying Principle to a formulation of type theory which differs slightly from \mathcal{Q}_0 is presented in [Andrews, 1971].

5503 Weak Compactness Theorem. If every finite subset of the set \mathcal{S} of sentences has a general model, then \mathcal{S} has a general model.

Proof: By 5403, every finite subset of \mathcal{S} is consistent. Therefore \mathcal{S} is consistent (since every proof is finite), so \mathcal{S} has a general model by 5501. ■

REMARK. It is natural to consider the Strong Compactness conjecture that if every finite subset of a set \mathcal{S} of sentences has a standard model, then \mathcal{S} has a standard model. We shall show that this is false at the end of §72.

An important application of the Weak Compactness Theorem is the construction of nonstandard models of the real number system, which leads

to the subject known as nonstandard analysis. The basic techniques are the same as those used to apply the Compactness Theorem for first-order logic (2514) in §25.

REMARK. Our results about frugal models are somewhat disturbing. As we shall see in §60, we can add to \mathcal{Q}_0 an axiom of infinity (which we shall call Axiom 6) which is consistent with \mathcal{Q}_0 (Corollary 7204) and which is true only in models whose domains of individuals are infinite. If $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{J} \rangle$ is any such model, Cantor's Theorem (X5304) must be valid in it, so it would appear that \mathcal{D}_α must be uncountably infinite, and hence that any model of Axiom 6 is uncountable. However, Henkin's Theorem implies that {Axiom 6} must have a countable model. This seemingly paradoxical situation is called Skolem's Paradox. (Recall the discussion following 2513.) It can be resolved by looking carefully at Cantor's Theorem, i.e., $\sim \exists g_{\text{on}} \forall f_{\text{oi}} \exists j_i \cdot g_{\text{on}} j_i = f_{\text{oi}}$, and considering what it means in a model. The theorem says that there is no function $g \in \mathcal{D}_{\text{on}}$ from \mathcal{D}_i into \mathcal{D}_α which has every set $f_{\text{oi}} \in \mathcal{D}_\alpha$ in its range. The usual interpretation of the statement is that \mathcal{D}_α is bigger (in cardinality) than \mathcal{D}_i . However, this obviously cannot be true in an infinite countable model, since \mathcal{D}_i and \mathcal{D}_α are both countably infinite, so there must be a bijection g between them. Of course, g cannot be in \mathcal{D}_{on} , or Cantor's Theorem would be contradicted. Thus, the model must be nonstandard.

We shall explore these ideas more carefully in 5504–5506 below. However, let us first consider what Henkin's Theorem tells us about models of the theory of real numbers. One can introduce finitely many constants to denote the basic operations and relations in the field of real numbers, and postulates which state that the domain of individuals is the set of real numbers (for example), and if one really believes in the real numbers, and expresses the postulates carefully, one would expect that these postulates would constitute a consistent set of sentences. Therefore, by Henkin's Theorem, there must be a countable model for these postulates, in spite of the fact that the set of real numbers is uncountable. Of course, the model will turn out to be nonstandard. One can express in \mathcal{Q}_0 , and prove with the aid of postulates, the statement that there are uncountably many real numbers, but in a nonstandard model this statement turns out not to mean quite what one expected it to. Even if one could take as postulates the set of all sentences true in the standard model whose domain of individuals is the set of real numbers, the resulting system would still have a countable model, as shown by Henkin's Theorem. Note that there is no sentence of \mathcal{Q}_0 by which one can distinguish between this countable model and the standard model of the real numbers (i.e., there is no sentence true in one model, but false

in the other). Thus one who speaks the language \mathcal{Q}_0 cannot tell whether he lives in a standard or nonstandard world, even if he can answer all the questions he can ask.

Of course, one can abstractly consider a system with uncountably many individual constants, say one to denote each real number. However, even this system would have a general model in which the set \mathcal{D}_{α} of sets of real numbers would have only the cardinality of the continuum.

5504 Lemma. Let $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{T} \rangle$ be a frugal model for \mathcal{Q}_0 such that $\text{card}(\mathcal{D}_i) = \text{card}(\mathcal{L}(\mathcal{Q}_0))$. Then \mathcal{M} is nonstandard.

Proof: Suppose \mathcal{M} were standard. Then all subsets of \mathcal{D}_i would be in $\mathcal{D}_{\alpha i}$. By Cantor's Theorem (see X5304), \mathcal{D}_i has more subsets than members, so $\text{card}(\mathcal{D}_{\alpha i}) > \text{card}(\mathcal{D}_i) = \text{card}(\mathcal{L}(\mathcal{Q}_0)) \geq \text{card}(\mathcal{D}_{\alpha i})$ (by frugality). This contradiction shows that \mathcal{M} must be nonstandard. ■

5505 Corollary. Every infinite frugal model for a countable formulation of \mathcal{Q}_0 is nonstandard.

5506 Theorem. Let \mathcal{H} be any set of sentences of \mathcal{Q}_0 which has infinite models. Then \mathcal{H} has nonstandard models.

Proof: Let $\kappa = \text{card}(\mathcal{L}(\mathcal{Q}_0))$ and let \mathcal{Q}_0^+ be an expansion of \mathcal{Q}_0 obtained by adding to \mathcal{Q}_0 a set \mathcal{C} of new individual constants, where $\text{card}(\mathcal{C}) = \kappa$. Let \mathcal{S} be the set of all sentences of the form $\mathbf{c}_i \neq \mathbf{d}_i$, where \mathbf{c}_i and \mathbf{d}_i are distinct constants in \mathcal{C} . It is easy to see that any finite subset of $\mathcal{S} \cup \mathcal{H}$ has a model, since the finitely many sentences from \mathcal{S} can all be made true in any infinite model of \mathcal{H} by giving the constants from \mathcal{C} which occur in those sentences distinct interpretations. Therefore, by 5503, 5403, and 5501 $\mathcal{S} \cup \mathcal{H}$ has a frugal general model $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{T} \rangle$. Since all sentences of \mathcal{S} are true in \mathcal{M} , \mathcal{D}_i has κ distinct members, so \mathcal{M} is nonstandard by 5504. ■

Thus, there are always nonstandard models for infinite mathematical systems, but one has the option of ignoring them, and concentrating on standard models. From the point of view of standard models, however, the system \mathcal{Q}_0 is not complete, since there are sentences true in all standard models which are not theorems. For example, the Axiom of Choice is presumably true in all standard models (though it seems difficult to prove this without making assumptions so strong as to rob the proof of any significance), but is false in certain nonstandard models. (Such a nonstandard model is used in [Andrews, 1972b] to prove the independence of the Axiom of Choice.)

It would be desirable, therefore, to add new axioms to \mathcal{Q}_0 to obtain a strengthened system in which the theorems are precisely the wffs_o valid in all standard models. Unfortunately, this is impossible, unless we abandon the rather fundamental principle that there must be an effective procedure for deciding, of an arbitrary wff, whether or not it is an axiom. We shall see this in Theorem 7206 of Chapter 7.

EXERCISES

X5500. Prove:

$$\begin{aligned}
 & \forall p_{o\beta} [\exists x_\beta p_{o\beta} x_\beta \supset p_{o\beta} \bullet j_{\beta(o\beta)} p_{o\beta}] \\
 & \supset \forall f_{\alpha\beta} \forall g_{\alpha\beta} [f_{\alpha\beta} [j_{\beta(o\beta)} \bullet \lambda x_\beta \bullet f_{\alpha\beta} x_\beta \neq g_{\alpha\beta} x_\beta] \\
 & = g_{\alpha\beta} [j_{\beta(o\beta)} \bullet \lambda x_\beta \bullet f_{\alpha\beta} x_\beta \neq g_{\alpha\beta} x_\beta] \supset \bullet f_{\alpha\beta} = g_{\alpha\beta}.
 \end{aligned}$$

X5501. Justify the Remark following the definition of extensional completeness.

X5502. A set of sentences is *categorical* iff all its models are isomorphic to each other. Prove that every categorical set of sentences is complete.

X5503. Is there a sentence **B** such that for every general model \mathcal{M} , it is true that $\mathcal{M} \models \mathbf{B}$ if and only if \mathcal{M} is a standard model?

Chapter 6

Formalized Number Theory

§60. Cardinal Numbers and the Axiom of Infinity

In the preface to *Principia Mathematica*, it is stated that “what were formerly taken, tacitly or explicitly, as axioms, are either unnecessary or demonstrable”.¹ In this spirit, we shall in this chapter define the natural numbers as the finite cardinals, and derive Peano’s Postulates from an Axiom of Infinity. We then establish those results about natural numbers and recursive functions which we need in order to establish the incompleteness results in Chapter 7.

We shall not develop the general theory of cardinal numbers, but it may be helpful to introduce and discuss a few relevant definitions.

DEFINITIONS. $E_{o(o\alpha)(o\beta)}$ stands for

$$[\lambda p_{o\beta} \lambda q_{o\alpha} \exists s_{\alpha\beta} \cdot \forall x_\beta [p_{o\beta} x_\beta \supset q_{o\alpha} \cdot s_{\alpha\beta} x_\beta] \\ \wedge \forall y_\alpha \cdot q_{o\alpha} y_\alpha \supset \exists_1 x_\beta \cdot p_{o\beta} x_\beta \wedge y_\alpha = s_{\alpha\beta} x_\beta].$$

$[NC]_{o(o(o\beta))}$ stands for $[\lambda u_{o(o\beta)} \exists p_{o\beta} \cdot u_{o(o\beta)} = E_{o(o\beta)(o\beta)} p_{o\beta}]$.

$E_{o(o\alpha)(o\beta)} p_{o\beta} q_{o\alpha}$ means that the sets $p_{o\beta}$ and $q_{o\alpha}$ are equipollent, i.e., $p_{o\beta}$ and $q_{o\alpha}$ have the same cardinality (number of members). This is expressed by saying that there is a function $s_{\alpha\beta}$ which is one-to-one on $p_{o\beta}$ and maps $p_{o\beta}$ onto $q_{o\alpha}$. It is easy to establish that $E_{o(o\beta)(o\beta)}$ is an equivalence relation.

¹[Whitehead and Russell, 1913, Second Edition, Volume 1, p. v.]

(See Exercises X6000-X6002.) The *cardinality* of a set $p_{\alpha\beta}$ (sometimes denoted $\overline{p_{\alpha\beta}}$) is $E_{\alpha(o\beta)(o\beta)}p_{\alpha\beta}$, the equivalence class of sets (of the same type as $p_{\alpha\beta}$) which are equipollent to $p_{\alpha\beta}$. $[NC]_{\alpha(o(o\beta))}$ is the set of *cardinal numbers* (for type β), i.e., the set of cardinalities (of sets of type $(o\beta)$).

In order to assure that all of the finite cardinals exist and are distinct, we must assume an axiom of infinity (i.e., a sentence which is satisfiable, but has no finite models.) The assertion that there are infinitely many individuals can be expressed naturally in a variety of ways, and in the absence of the Axiom of Choice it is not always possible to prove the equivalence of various natural formulations of this assertion. For our Axiom of Infinity we shall take the assertion that there is an irreflexive partial ordering of the individuals with respect to which there is no maximal element.²

Axiom 6:

$$\exists r_{on} \forall x_i \forall y_i \forall z_i [\exists w_i r_{on} x_i w_i \wedge \sim r_{on} x_i x_i \\ \wedge \neg r_{on} x_i y_i \supset \neg r_{on} y_i z_i \supset r_{on} x_i z_i].$$

Without going into the subject very far, let us consider some alternative axioms of infinity which might naturally occur to the reader. One way to say that a set is infinite is to say that it is equipollent with a proper subset of itself. (For example, the set of integers is equipollent with the set of even integers.) This can be expressed by the wff

$$Inf^I : \exists n_{oi} \exists m_{oi} [\forall x_i [m_{oi} x_i \supset n_{oi} x_i] \\ \wedge \exists o_i [n_{oi} o_i \wedge \sim m_{oi} o_i] \wedge E_{o(oi)(oi)} n_{oi} m_{oi}].$$

Another way to say that a set is infinite is to say that it has a subset which is isomorphic to the set of natural numbers. The natural numbers can be characterized by Peano's Postulates, which can be stated informally as follows:

- (P1) There is an entity called 0 which is a natural number.
- (P2) Every natural number n has a successor S_n which is also a natural number.
- (P3) 0 is not the successor of any natural number.

²This axiom of infinity was first suggested (in essence) in [Bernays and Schönfinkel, 1928].

- (P4) If n and m are natural numbers with the same successors, then n and m are the same.
 (P5) Principle of Mathematical Induction (see §10A).

Actually, any set satisfying postulates P1–P4 must be infinite. The Induction Principle simply limits the size of the set. Postulates P1–P4 can be expressed by the wff

$$\begin{aligned} \text{Inf}^{II} : \quad & \exists n_{oi} \exists o_i \exists s_{ii} [n_{oi} o_i \wedge \forall x_i [n_{oi} x_i \supset n_{oi} = s_{ii} x_i] \\ & \wedge \forall x_i [n_{oi} x_i \supset s_{ii} x_i \neq o_i] \\ & \wedge \forall x_i \forall y_i [n_{oi} x_i \wedge n_{oi} y_i \wedge s_{ii} x_i = s_{ii} y_i \supset x_i = y_i]. \end{aligned}$$

It is not hard to see that $\vdash_{Q_0} \text{Inf}^I \equiv \text{Inf}^{II}$. Indeed, if we assume that there are sets n_{oi} and m_{oi} of individuals and there is an individual o_i as described in Inf^I , and let s_{ii} be a function demonstrating that $\text{En}_{oi} m_{oi}$, then n_{oi} , o_i , and s_{ii} verify Inf^{II} , so $\vdash \text{Inf}^I \supset \text{Inf}^{II}$. On the other hand, if there exist n_{oi} , o_i , and s_{ii} as described in Inf^{II} , and m_{oi} is the image of n_{oi} under s_{ii} , then n_{oi} and m_{oi} verify Inf^I , so $\vdash \text{Inf}^{II} \supset \text{Inf}^I$.

It can also be seen that $\vdash_{Q_0} \text{Inf}^{II} \supset \text{Axiom } 6$. The details of this argument are best postponed to §62, but we may give a preview of it here. If there exist n_{oi} , o_i , and s_{ii} as described in Inf^{II} , we may let n'_{oi} be the intersection of all sets containing o_i and closed under s_{ii} (as in the definition of \mathbb{N}_{oi} below). Then n'_{oi} is isomorphic to the natural numbers, and the natural ordering $<$ is definable on n'_{oi} (as in §62). The relation $[\lambda x_i \lambda y_i [n'_{oi} y_i \wedge \sim n'_{oi} x_i \vee x_i < y_i]]$ satisfies the requirements on r_{oi} in Axiom 6.

Thus, all the consequences of Axiom 6 which we shall derive below are also consequences of Inf^I and Inf^{II} , so any of these three sentences could serve as our axiom of infinity. Of course, there are many additional possibilities.

We shall use Q_0^∞ as a name for the system obtained by adding Axiom 6 to the list of axioms of Q_0 . Henceforth in this chapter, $\vdash \mathbf{A}$ shall mean that \mathbf{A} is a theorem of Q_0^∞ (i.e., $\vdash_{Q_0^\infty} \mathbf{A}$) unless otherwise indicated. Note that $\vdash_{Q_0^\infty} \mathbf{A}$ iff [Axiom 6] $\vdash_{Q_0} \mathbf{A}$, and that Axiom 6 is a sentence, so we can use the derived rules of inference established in §52 for Q_0^∞ as well as for Q_0 . In presenting (condensed) proofs, we shall often omit explicit reference to the rules of inference used to infer a line of the proof, and simply indicate from what lines of the proof and previously established theorems the given line is

obtained. When a line of a proof is inferred from the immediately preceding line, we may also omit explicit mention of the preceding line. From time to time the reader may find it a useful exercise to fill in all intermediate steps of a proof, and indicate the rule of inference used for each step. We shall often omit type symbols from occurrences of a variable after the first occurrence in a wff, or when the context in some other way makes it clear what the type must be.

We next define the natural numbers. These are equivalence classes of sets of individuals, and so have type $(o(o))$.

DEFINITIONS. Let σ be the type symbol $(o(o))$.

0_σ stands for $Q_{o(o)}(o)[\lambda x_i F_o]$

$S_{\sigma\sigma}$ stands for $[\lambda n_{o(o)} \lambda p_{o\sigma} \exists x_i \cdot p_{o\sigma} x_i \wedge n_{o(o)} [\lambda t_i \cdot t_i \neq x_i \wedge p_{o\sigma} t_i]]$

$\mathbb{N}_{o\sigma}$ stands for $[\lambda n_\sigma \forall p_{o\sigma} \cdot [p_{o\sigma} 0_\sigma \wedge \forall x_\sigma \cdot p_{o\sigma} x_\sigma \supset p_{o\sigma} S_{\sigma\sigma} x_\sigma] \supset p_{o\sigma} n_\sigma]$

$Finite_{o(o)}$ stands for $[\lambda p_{o\sigma} \exists n_\sigma \cdot \mathbb{N}_{o\sigma} n_\sigma \wedge n_\sigma p_{o\sigma}]$

Thus zero is the collection of all sets with zero members, i.e., the collection containing just the empty set $[\lambda x_i F_o]$. The wff S represents the successor function. If $n_{(o(o))}$ is a finite cardinal (say 2), then a set $p_{o\sigma}$ (say $\{a, b, c\}$) is in Sn iff there is an individual (say c) which is in $p_{o\sigma}$ and whose deletion from $p_{o\sigma}$ leaves a set ($\{a, b\}$) which is in n . $\mathbb{N}_{o\sigma}$ represents the set of natural numbers, i.e., the intersection of all sets which contain 0 and are closed under S . (See Exercise X6007.) It can be shown that every natural number is a cardinal number (Exercise X6102).

DEFINITIONS. Let $M_{o\alpha}$ be any wff in which x_α does not occur free.

$(\forall x_\alpha \in M_{o\alpha}) A$ stands for $\forall x_\alpha [M_{o\alpha} x_\alpha \supset A]$.

$(\exists x_\alpha \in M_{o\alpha}) A$ stands for $\exists x_\alpha [M_{o\alpha} x_\alpha \wedge A]$.

$\dot{\forall} x_\sigma A$ stands for $(\forall x_\sigma \in \mathbb{N}_{o\sigma}) A$.

$\dot{\exists} x_\sigma A$ stands for $(\exists x_\sigma \in \mathbb{N}_{o\sigma}) A$.

We leave to the reader the proofs of the following simple but useful propositions of Q_0 :

6000. $\vdash \sim (\forall x_\alpha \in M_{o\alpha}) A = (\exists x_\alpha \in M_{o\alpha}) \sim A$ provided x_α is not free in $M_{o\alpha}$.

6001. $\vdash \sim (\exists x_\alpha \in M_{o\alpha}) A = (\forall x_\alpha \in M_{o\alpha}) \sim A$ provided x_α is not free in $M_{o\alpha}$.

EXERCISES

Prove the following theorems:

$$\mathbf{X6000. } \forall p_{o\beta} \bullet E_{o(o\beta)(o\beta)} p_{o\beta} p_{o\beta}$$

$$\mathbf{X6001. } \forall p_{o\beta} \forall q_{o\alpha} \bullet E_{o(o\alpha)(o\beta)} p_{o\beta} q_{o\alpha} \supset E_{o(o\beta)(o\alpha)} q_{o\alpha} p_{o\beta}$$

$$\mathbf{X6002. } \forall p_{o\beta} \forall q_{o\alpha} \forall r_{o\gamma} \bullet E_{o(o\alpha)(o\beta)} p_{o\beta} q_{o\alpha} \wedge E_{o(o\gamma)(o\alpha)} q_{o\alpha} r_{o\gamma} \supset E_{o(o\gamma)(o\beta)} p_{o\beta} r_{o\gamma}$$

$$\mathbf{X6003. } \forall p_{o\beta} \forall q_{o\beta} \bullet E_{o(o\beta)(o\beta)} p_{o\beta} q_{o\beta} = \bullet \overline{\overline{p_{o\beta}}} = \overline{\overline{q_{o\beta}}}$$

$$\mathbf{X6004. } E_{o(o\alpha)(o\beta)} [Q_{o\beta\beta} x_\beta] [Q_{o\alpha\alpha} y_\alpha]$$

$$\mathbf{X6005. } [NC]_{o(o(o\beta))} \overline{\overline{p_{o\beta}}}$$

$$\mathbf{X6006. } Inf^I = Inf^{II}$$

$$\mathbf{X6007. } \mathbb{N} = \bigcap_{o(o\sigma)} [\lambda p_{o\sigma} \bullet p_{o\sigma} 0_\sigma \wedge \forall x_\sigma \bullet p_{o\sigma} x_\sigma \supset p_{o\alpha} \bullet S_{\sigma\sigma} x_\sigma]$$

X6008. Is there a sentence S such that for every general model \mathcal{M} , it is true that $\mathcal{M} \models S$ if and only if \mathcal{M} has an infinite domain of individuals?
(Hint: use 5503.)

§61. Peano's Postulates

We now prove Peano's Postulates (Theorems 6100, 6101, 6102, 6106, and 6109), plus a few related results.

$$\mathbf{6100. } \vdash \mathbb{N}_{o\sigma} 0_\sigma \quad \text{by the def of } \mathbb{N}$$

$$\mathbf{6101. } \vdash \forall x_\sigma \bullet \mathbb{N}_{o\sigma} x_\sigma \supset \mathbb{N}_{o\sigma} \bullet S_{o\sigma} x_\sigma$$

Proof:

- (.1) $\mathbb{N} x_\sigma, .1 \vdash p_{o\alpha} 0 \wedge \forall x_\sigma \bullet px \supset p \bullet Sx$ hyp
- (.2) $\mathbb{N} x_\sigma, .1 \vdash p_{o\sigma} x_\sigma$.1, hyp., def of \mathbb{N}
- (.3) $\mathbb{N} x_\sigma, .1 \vdash p_{o\sigma} \bullet Sx_\sigma$.1, .2
- (.4) $\mathbb{N} x_\sigma \vdash \mathbb{N} \bullet Sx_\sigma$.3, def of \mathbb{N}

■

6102. The Induction Theorem.

$$\vdash \forall p_{o\sigma} \cdot [p_{o\sigma} 0_\sigma \wedge \dot{\forall} x_\sigma \cdot p_{o\sigma} x_\sigma \supset p_{o\sigma} \cdot S_{\sigma\sigma} x_\sigma] \supset \dot{\forall} x_\sigma p_{o\sigma} x_\sigma$$

Proof: Let $P_{o\sigma}$ be $[\lambda t_\sigma \cdot \mathbb{N} t \wedge p_{o\sigma} t]$.

- (.1) $.1 \vdash p_{o\sigma} 0 \wedge \forall x_\sigma \cdot \mathbb{N} x \supset \cdot px \supset p \cdot Sx$ hyp
- (.2) $\mathbb{N} y_\sigma \vdash [P_0 \wedge \forall x_\sigma \cdot Px \supset P \cdot Sx] \supset Py_\sigma$ hyp, def of \mathbb{N}
- (.3) $.1 \vdash P_0$ def of P , .1, 6100
- (.4) $.1 \vdash \forall x_\sigma \cdot Px \supset P \cdot Sx$ def of P , .1, 6101
- (.5) $.1, \mathbb{N} y_\sigma \vdash Py_\sigma$.2, .3, .4
- (.6) $.1 \vdash \dot{\forall} y_\sigma py_\sigma$.5, def of $\dot{\forall}$, P

■

$$\mathbf{6103.} \vdash p_{o\sigma} 0_\sigma \wedge \dot{\forall} x_\sigma p_{o\sigma} S_{\sigma\sigma} x_\sigma \supset \dot{\forall} n_\sigma p_{o\sigma} n_\sigma$$

Proof:

- (.1) $\vdash \dot{\forall} x_\sigma p_{o\sigma} S_{\sigma\sigma} x_\sigma \supset \dot{\forall} x_\sigma \cdot p_{o\sigma} x_\sigma \supset p_{o\sigma} S_{\sigma\sigma} x_\sigma$ Rule Q
- The theorem follows by 6102. ■

$$\mathbf{6104.} \vdash \dot{\forall} n_\sigma \cdot n_\sigma = 0 \vee \dot{\exists} m_\sigma \cdot n_\sigma = S_{\sigma\sigma} m_\sigma$$

Proof by 6103:

- (.1) $\vdash 0_\sigma = 0_\sigma \vee \dot{\exists} m_\sigma \cdot 0_\sigma = S_{\sigma\sigma} m_\sigma$ 5200
- (.2) $\mathbb{N} x_\sigma \vdash \mathbb{N} x_\sigma \wedge \cdot Sx_\sigma = Sx_\sigma$ 5200
- (.3) $\mathbb{N} x_\sigma \vdash \dot{\exists} m_\sigma \cdot \mathbb{N} m_\sigma \wedge Sx_\sigma = Sm_\sigma$ \exists Gen: .2
- (.4) $\vdash \forall x_\sigma \cdot Sx_\sigma = 0 \vee \dot{\exists} m_\sigma \cdot Sx_\sigma = Sm_\sigma$.3

Then use 6103 with .1 and .4. ■

$$\mathbf{6105.} \vdash \forall n_\sigma \cdot S_{\sigma\sigma} n_\sigma \neq 0_\sigma$$

Proof by contradiction:

- (.1) $.1 \vdash Sn_\sigma = 0$ hyp
- (.2) $\vdash 0_\sigma [\lambda x_i F]$ def of 0
- (.3) $.1 \vdash Sn_\sigma [\lambda x_i F]$ $=: .1, .2$
- (.4) $.1 \vdash \dot{\exists} x_i F$.3, def of S
- (.5) $\vdash Sn_\sigma \neq 0$.4

■

6106. $\vdash \dot{\forall} n_\sigma \bullet S_{\sigma\sigma} n_\sigma \neq 0_\sigma$ by 6105 and the def of $\dot{\forall}$

We have now proven all of Peano's Postulates except one, which will be proved below as Theorem 6109. Our first step in proving it is to show that if we remove any element from a set of cardinality S_n we obtain a set of cardinality n .

6107. $\vdash \dot{\forall} n_\sigma \forall p_{\sigma i} \bullet S_{n_\sigma} p_{\sigma i} \wedge p_{\sigma i} w_i \supset n_\sigma [\lambda t_i \bullet t_i \neq w_i \wedge p_{\sigma i} t_i]$

Proof: The proof is by induction on n . First we treat the case $n = 0$.

- (.1) .1 $\vdash S_0 p_{\sigma i} \wedge p w_i$ hyp
- (.2) .1 $\vdash \exists x_i . 3$.1, def of S
- (.3) .1, .3 $\vdash p_{\sigma i} x_i \wedge 0[\lambda t_i \bullet t \neq x_i \wedge p t]$ choose x
- (.4) .1, .3 $\vdash \sim [\lambda t_i \bullet t \neq x_i \wedge p_{\sigma i} t] w_i$.3, def of 0
- (.5) .1, .3 $\vdash w_i = x_i$.1, .4
- (.6) .1, .3 $\vdash 0[\lambda t_i \bullet t \neq w_i \wedge p_{\sigma i} t]$.3, .5
- (.7) .1 $\vdash .6$.2, .6
- (.8) $\vdash \forall p_{\sigma i} \bullet S_0 p \wedge p w_i \supset 0[\lambda t_i \bullet t \neq w_i \wedge p t]$.7

Next we treat the induction step.

- (.9) .9 $\vdash \mathbb{N} n_\sigma \wedge \forall p_{\sigma i} \bullet S_{n_\sigma} p \wedge p w_i \supset n[\lambda t_i \bullet t \neq w_i \wedge p t]$ (inductive) hyp
- (.10) .10 $\vdash [S S n_\sigma] p_{\sigma i} \wedge p w_i$ hyp
- (.11) .10 $\vdash \exists x_i . 12$.10, def of S
- (.12) .10, .12 $\vdash p_{\sigma i} x_i \wedge S_{n_\sigma} [\lambda t_i \bullet t \neq x_i \wedge p t]$ choose x

From .10 we must prove $S_n [\lambda t \bullet t \neq w_i \wedge p t]$. We consider two cases in .13 and .15.

- (.13) .13 $\vdash x_i = w_i$ hyp (case 1)
- (.14) .10, .12, .13 $\vdash S_{n_\sigma} [\lambda t_i \bullet t \neq w_i \wedge p_{\sigma i} t]$.12, .13

In case 2 we shall use the inductive hypothesis.

- (.15) .15 $\vdash x_i \neq w_i$ hyp (case 2)
- (.16) .9 $\vdash S_{n_\sigma} [\lambda t_i \bullet t \neq x_i \wedge p_{\sigma i} t]$
 - $\wedge [\lambda t_i \bullet t \neq x_i \wedge p t] w_i \supset n[\lambda t_i \bullet t \neq w_i \wedge t \neq x_i \wedge p t]$.9
- (.17) .9, .10, .12, .15 $\vdash n_\sigma [\lambda t_i \bullet t \neq w_i \wedge t \neq x_i \wedge p_{\sigma i} t]$.10, .12, .15, .16
- (.18) .9, .10, .12, .15 $\vdash x_i \neq w_i \wedge p_{\sigma i} x_i \wedge n_\sigma [\lambda t_i \bullet t \neq x_i \wedge t \neq w_i \wedge p t]$.12, .15, .17
- (.19) .9, .10, .12, .15 $\vdash S_{n_\sigma} [\lambda t_i \bullet t \neq w_i \wedge p_{\sigma i} t]$.18, def of S

This completes case 2.

- (.20) .9, .10 $\vdash Sn_{\sigma}[\lambda t_i \cdot t \neq w_i \wedge p_{oi}t]$.14, .19, .11
(.21) .9 $\vdash \forall p_{oi} \cdot [SSn_{\sigma}]p \wedge pw_i \supset Sn[\lambda t_i \cdot t \neq w \wedge pt]$.20

This completes the induction step. The theorem now follows from .8 and .21 by the Induction Theorem. ■

REMARK. It can now be seen (Exercise X6107) that

$$\vdash S_{\sigma\sigma}n_{\sigma}p_{oi} = \square \sim 0_{\sigma}p \wedge \forall x_i \cdot px \supset n_{\sigma}[\lambda t_i \cdot t \neq x \wedge pt].$$

This provides an alternative possible definition for $S_{\sigma\sigma}$.

It will be observed so far in this section we have not used the Axiom of Infinity. We shall use it in proving the next theorem, which will be used to prove Theorem 6109.

$$\mathbf{6108.} \vdash \dot{\forall}n_{\sigma} \cdot n_{\sigma}p_{oi} \supset \exists w_i \sim p_{oi}w_i$$

Proof:

- (.1) .1 $\vdash \forall x_i \forall y_i \forall z_i \cdot \exists w_i r_{oi}xw \wedge \sim rxx \wedge \square rxy \supset \square ryx \supset rxz$
choose r (Axiom 6)

Let P_{oi} be $[\lambda n_{\sigma} \forall p_{oi} \cdot np \supset \exists z_i \forall w_i \cdot pw \supset \sim r_{oi}zw]$. We may informally interpret rzw as meaning that z is below w . Thus P_n means that if p is any set of cardinality n , then there is an element z which is below no member of p . We shall prove $\dot{\forall}n_{\sigma}P_n$ by induction on n .

- (.2) $0p_{oi} \vdash \sim p_{oi}w_i$ def of 0
(.3) $\vdash P_0$.2, def of P

Next we treat the induction step.

- (.4) .4 $\vdash \mathbb{N}n_{\sigma} \wedge P_n$ (inductive) hyp
(.5) .5 $\vdash Sn_{\sigma}p_{oi}$ hyp
(.6) .5 $\vdash \exists x_i . 7$.5, def of S
(.7) .5, .7 $\vdash p_{oi}x_i \wedge n_{\sigma}[\lambda t_i \cdot t \neq x_i \wedge pt]$ choose x (.6)
(.8) .4, .5, .7 $\vdash \exists z_i . 9$.4, def of P , .7
(instantiate $\forall p_{oi}$ in P_n with $[\lambda t_i \cdot t \neq x_i \wedge p_{oi}t]$)
(.9) .4, .5, .7 .9 $\vdash \forall w_i \cdot w \neq x_i \wedge p_{oi}w \supset \sim r_{oi}zw$ choose z (.8)

Thus from the inductive hypothesis we see that there is an element z which is below nothing in $p - \{x\}$. We must show that there is an element which is below nothing in p . We consider two cases, .10 and .14.

$$(10) .10 \vdash \sim r_{on}zx \quad \text{hyp (case 1)}$$

In this case we show that z is under nothing in p .

$$(11) .4, .5, .7, .10 \vdash w_i = x_i \supset \sim r_{on}zw \quad .10, \text{Axiom 2}$$

$$(12) .4, .5, .7, .9, .10 \vdash \forall w_i \bullet p_{oi}w \supset \sim r_{on}zw \quad .9, .11$$

$$(13) .4, .5, .7, .9, .10 \vdash \exists z_i .12 \quad .12$$

Next we consider case 2, and show by the indirect method that x is below nothing in p .

$$\begin{aligned} (14) .14 \vdash r_{on}zx & \quad \text{hyp (case 2)} \\ (15) .1, .14, .15 \vdash p_{oi}w_i \wedge r_{on}x_iw & \quad \text{hyp} \\ (16) .1, .14, .15 \vdash r_{on}z_iw_i & \quad .1, .14, .15 \\ (17) .1, .9, .14, .15 \vdash w_i = x_i & \quad .9, .15, .16 \\ (18) .1, .9, .14, .15 \vdash r_{on}x_ix_i & \quad .15, .17 \\ (19) .1, .4, .5, .7, .9, .14, .15 \vdash F_o & \quad .1, .18 \\ (20) .1, .4, .5, .7, .9, .14 \vdash \forall w_i \bullet p_{oi}w \supset \sim r_{on}x_iw & \quad .19 \\ (21) .1, .4, .5, .7, .9, .14 \vdash \exists z_i \forall w_i \bullet p_{oi}w \supset \sim r_{on}zw & \quad .20 \\ (22) .1, .4, .5 \vdash .21 & \quad \text{Cases: .13, .21; Rule C: .8, .6} \\ (23) .1 \vdash \dot{\mathbb{N}}n_\sigma \wedge Pn \supset P \bullet Sn & \quad .22, \text{def of } P \\ (24) .1 \vdash \dot{\mathbb{N}}n_\sigma Pn_\sigma & \quad .3, .23, \text{Induction Theorem} \end{aligned}$$

Having finished the inductive proof, we proceed to prove the main theorem.

$$\begin{aligned} (25) .25 \vdash \dot{\mathbb{N}}n_\sigma \wedge np_{oi} & \quad \text{hyp} \\ (26) .1, .25 \vdash \exists z_i \forall w_i \bullet p_{oi}w \supset \sim r_{on}zw & \quad .24, .25, \text{def of } P \\ (27) .1 \vdash \forall z_i \exists w_i r_{on}zw & \quad .1 \\ (28) .1, .25 \vdash \exists w_i \sim p_{oi}w & \quad .26, .27 \\ (29) .1 \vdash \dot{\mathbb{N}}n_\sigma \bullet np_{oi} \supset \exists w_i \sim p_{oi}w_i & \quad .28 \\ (30) \vdash .29 & \quad \text{Rule C: Axiom 6, .29} \end{aligned}$$

■

6109. $\vdash \forall n_\sigma \forall m_\sigma \bullet S_{\sigma\sigma} n_\sigma = S_{\sigma\sigma} m_\sigma \supset n_\sigma = m_\sigma$

Proof:

- (.1) $.1 \vdash \mathbb{N} n_\sigma \wedge \mathbb{N} m_\sigma \wedge Sn = Sm$ hyp
- (.2) $.2 \vdash n_\sigma p_{\sigma i}$ hyp
- (.3) $.1, .2 \vdash \exists w_i \sim p_{\sigma i} w$.1, .2, 6108
- (.4) $.1, .2, .4 \vdash \sim p_{\sigma i} w_i$ choose w (.3)
- (.5) $.1, .2, .4 \vdash p_{\sigma i} = [\lambda t_i \bullet t \neq w_i \wedge \bullet t = w \vee p_{\sigma i} t]$.4, Axiom 2, Axiom 3
- (.6) $.1, .2, .4 \vdash n_\sigma [\lambda t_i \bullet t \neq w_i \wedge \bullet t = w \vee p_{\sigma i} t]$.2, .5
- (.7) $.1, .2, .4 \vdash Sn_\sigma [\lambda t_i \bullet t = w_i \vee p_{\sigma i} t]$.6, def of S
- (.8) $.1, .2, .4 \vdash Sm_\sigma [\lambda t_i \bullet t = w_i \vee p_{\sigma i} t]$.1, .7
- (.9) $.1, .2, .4 \vdash Sm_\sigma [\lambda t_i \bullet t = w_i \vee p_{\sigma i} t] \wedge [\lambda t_i \bullet t = w \vee p_{\sigma i} t] w$
 $\supset m[\lambda t_i \bullet t \neq w \wedge \bullet t = w \vee p_{\sigma i} t]$.1, 6107
- (.10) $.1, .2, .4 \vdash m_\sigma [\lambda t_i \bullet t \neq w_i \wedge \bullet t = w \vee p_{\sigma i} t]$.8, .9, 5200
- (.11) $.1, .2, .4 \vdash m_\sigma p_{\sigma i}$.5, .10
- (.12) $.1 \vdash n_\sigma p_{\sigma i} \supset m_\sigma p$.3, .11
- (.13) $.1 \vdash m_\sigma p_{\sigma i} \supset n_\sigma p$ Deduction Theorem, Sub: .12
- (.14) $.1 \vdash \forall p_{\sigma i} \bullet n_\sigma p = m_\sigma p$.12, .13
- (.15) $.1 \vdash n_\sigma = m_\sigma$.14, Axiom 3

■

This completes the derivation of Peano's Postulates for the natural numbers. Note that the Axiom of Infinity was needed to prove only the last of these. Henceforth our development of number theory will be based entirely on Peano's Postulates without further reference to the Axiom of Infinity or to the definitions of the natural numbers. Indeed, the reader may henceforth imagine that we are developing a language \mathcal{Q}^σ which is like \mathcal{Q}_0^∞ except that the type of individuals is called σ rather than i . Also, \mathcal{Q}^σ has primitive constants 0_σ and $S_{\sigma\sigma}$, and has 6105 and 6109 as axioms in place of Axiom 6. $\mathbb{N}_{\sigma\sigma}$ is defined as above, and 6100–6104 are theorems of \mathcal{Q}^σ .

DEFINITIONS. Given wffs $F_{\alpha\alpha}$ and A_α , we define the wff $_n F_{\alpha\alpha}^n A_\alpha$ for each natural number $n \geq 0$ as follows by induction on n :

- (a) $F_{\alpha\alpha}^0 A_\alpha$ is A_α .
- (b) $F_{\alpha\alpha}^{n+1} A_\alpha$ is $[F_{\alpha\alpha} \bullet F_{\alpha\alpha}^n A_\alpha]$.

If n is any natural number, \bar{n} is $[S_{\sigma\sigma}^n 0_\sigma]$; this wff is called the *representation* of n in \mathcal{Q}^∞ . The wffs \bar{n} represent natural numbers and are called *numerals*.

We also wish to define what it means for a wff to represent a function of natural numbers. This could be expressed by saying that the wff denotes the function in some specified model(s), but we wish to have a purely syntactic definition of representability, so we use the following:

DEFINITION. Let $k \geq 1$, and let g be a k -ary function from natural numbers to natural numbers. The cwff $\mathbf{G}_{\sigma\sigma\dots\sigma}$ represents g in any extension of \mathcal{Q}_0^∞ iff for all k -tuples (n_1, \dots, n_k) of natural numbers,

$$\vdash \mathbf{G}_{\sigma\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k = \overline{g(n_1, \dots, n_k)}.$$

Let $k \geq 1$, and let g be a k -ary relation between natural numbers. (If $k = 1$, let g be a set of natural numbers.) The cwff $\mathbf{G}_{o\sigma\dots\sigma}$ represents g iff for all k -tuples (n_1, \dots, n_k) of natural numbers,

$$\begin{aligned} \vdash \mathbf{G}_{o\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k &= T_o && \text{if } g(n_1, \dots, n_k), \text{ and} \\ \vdash \mathbf{G}_{o\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k &= F_o && \text{if not } g(n_1, \dots, n_k). \end{aligned}$$

A numerical function or relation or set is *representable* in \mathcal{Q}_o^∞ iff there is a wff which represents it.

6110 $\vdash \mathbb{N}_{o\sigma} \bar{n}$ where n is any natural number.

The proof is by induction (in the meta-language) on n , using 6100 and 6101.

6111. If k is a natural number greater than 0, $\vdash \dot{\forall} x_\sigma \bullet S^k x_\sigma \neq x_\sigma$.

Proof: By induction on x .

- (.1) $\vdash S_{\sigma\sigma}[S^{k-1}0_\sigma] \neq 0_\sigma$ 6105
- (.2) $\mathbb{N}x_\sigma \vdash \mathbb{N}S^k x_\sigma$ by 6101 (used k times)
- (.3) $\mathbb{N}x_\sigma \vdash S_{\sigma\sigma}[S^k x_\sigma] = S_{\sigma\sigma}x_\sigma \supset \bullet S^k x_\sigma = x_\sigma$ 6109, .2
- (.4) $\vdash \forall x_\sigma \bullet S^k x_\sigma \neq x_\sigma \supset \bullet S^k[S_{\sigma\sigma}x_\sigma] \neq S_{\sigma\sigma}x_\sigma$.3, def of $S^k \mathbf{A}_\sigma$

Then use 6102 with .1 and .4. ■

6112 Theorem. $\mathcal{Q}_{o\sigma\sigma}$ represents the equality relation. Thus, if m and n are distinct natural numbers, $\vdash \bar{m} \neq \bar{n}$.

Proof: If $m = n$, then $\vdash \bar{m} = \bar{n}$ by 5200. If $m \neq n$, we may assume $m > n$, and let $k = m - n$.

$$(1) \vdash \mathbb{N}\bar{n} \supset \bullet S^k \bar{n} \neq \bar{n}$$

6111

(2) $\vdash \bar{m} \neq \bar{n}$ 6110, .1, def of \bar{m}

■

DEFINITION. The *characteristic function* χ_r of a k -ary relation r is defined as follows:

$$\begin{aligned}\chi_r(n_1, \dots, n_k) &= 1 && \text{if } r(n_1, \dots, n_k) \\ &= 0 && \text{if not } r(n_1, \dots, n_k).\end{aligned}$$

6113 Theorem. Let r be any relation between natural numbers. Then r is representable iff χ_r is representable.

Proof: Let k be the number of arguments of r . If r is represented by $G_{o\sigma\dots\sigma}$, let $C_{o\sigma\sigma\sigma}$ be the wff of theorem 5313 and let $H_{\sigma\sigma\dots\sigma}$ be $[\lambda x_\sigma^1 \dots \lambda x_\sigma^k \cdot C_{o\sigma\sigma\sigma} \bar{1} \bar{0} \cdot G_{o\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k]$. We show that $H_{\sigma\sigma\dots\sigma}$ represents χ_r . Let n_1, \dots, n_k be natural numbers.

$$(1) \vdash H_{\sigma\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k = C_{o\sigma\sigma\sigma} \bar{1} \bar{0} \cdot G_{o\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k \quad \text{def of } H$$

If $r(n_1, \dots, n_k)$, then

$$(2) \vdash G_{o\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k = T \quad \text{since } G \text{ represents } r$$

$$(3) \vdash H_{\sigma\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k = \bar{1} \quad .1, .2, 5313$$

If not $r(n_1, \dots, n_k)$, then

$$(4) \vdash G_{o\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k = F \quad \text{since } G \text{ represents } r$$

$$(5) \vdash H_{\sigma\sigma\dots\sigma} \bar{n}_1 \dots \bar{n}_k = \bar{0} \quad .1, .4, 5313$$

Thus $H_{\sigma\sigma\dots\sigma}$ does indeed represent χ_r .

On the other hand, if a wff $H_{\sigma\sigma\dots\sigma}$ represents χ_r , it is easy to see with the aid of 5200 and 6112 that $[\lambda x_\sigma^1 \dots \lambda x_\sigma^k \cdot Q_{o\sigma\sigma} \bar{1} \cdot H_{\sigma\sigma\dots\sigma} x_\sigma^1 \dots x_\sigma^k]$ represents r . ■

EXERCISES

X6100. Let f be the function from natural numbers to natural numbers such that $f(0) = 3$ and $f(n+1) = f(n)+1$ for all natural numbers n . Show that f is representable in Q_0^∞ by giving an explicit, detailed, and complete demonstration that it satisfies the definition of representability.

Prove the theorems stated in exercises X6101-X6108.

X6101. $\bar{1} = \sum_{o(o)}^1$

X6102. $\forall n_\sigma \cdot \mathbb{N}_{o\sigma} n_\sigma \supset [NC]_{o\sigma} n_\sigma$ (Every natural number is a cardinal number).

X6103. $\dot{\forall} n_\sigma \sim n_\sigma [\lambda x_i T_o]$

X6104. The *composition* of functions f and g is the function $f \circ g$ defined as follows: $(f \circ g)(x) = f(g(x))$. An *iterate* of a function g is a function of the form $g \circ \dots \circ g$ (the composition of k copies of g , where k is any natural number). In particular, we regard the identity function (here $k = 0$), g itself ($k = 1$), and $g \circ g$ ($k = 2$) as iterates of g .

- (a) Define a closed wff $I_{o(aa)(aa)}$ such that $I_{o(aa)(aa)} g_{aa}$ is a name in type theory for the set of iterates of g_{aa} .
- (b) Prove $I_{o(aa)(aa)} g_{aa} [\lambda x_a x_a]$.
- (c) Prove $I_{o(aa)(aa)} g_{aa} [\lambda x_a \bullet g_{aa} \bullet g_{aa} x_a]$.
- (d) Prove $I_{o(aa)(aa)} [\lambda x_a y_a] f_{aa} \supset f_{aa} y_a = y_a$.

X6105. $\forall n_{o(o\iota)} \cdot \mathbb{N} n \supset \forall q_{o\iota} \cdot n q \supset \exists j_{i(o\iota)} \forall r_{o\iota} \cdot r \subseteq q \wedge \exists x_i \forall x \supset r. j r$

X6106. $Finite[\lambda x_i T_o] \supset AC^i$

X6107. $S_{\sigma\sigma} n_\sigma p_{o\iota} = \bullet \sim 0_\sigma p \wedge \forall x_i \bullet px \supset n_\sigma [\lambda t_i \bullet t \neq x \wedge pt]$

X6108. $\bar{n} p_{o\iota} \supset \exists x_i^1 \dots \exists x_i^n \bullet \bigwedge_{i=1}^n p_{o\iota} x_i^i \wedge \bigwedge_{1 \leq i < j \leq n} x_i^i \neq x_j^j$ for each positive integer n . (Hint: use induction.)

X6109. Show that if \mathcal{M} is a model for \mathcal{Q}_0 with exactly n individuals (members of \mathcal{D}_i), $\mathcal{M} \models \bar{n} = Q_{o(o\iota)(o\iota)} [\lambda x_i T_o]$ and $\mathcal{M} \models \bar{m} = [\lambda p_{o\iota} F_o]$ for each integer $m > n$.

X6110. Show that if n is any positive integer and \mathcal{M} is any model of \mathcal{Q}_0 , then $\mathcal{M} \models \exists p_{o\iota} \bar{n} p_{o\iota}$ if and only if \mathcal{M} has at least n individuals.

§62. Order

Intuitively, if x and y are natural numbers, $x \leq y$ iff one can get from x to y by zero or more applications of the successor operation, i.e., iff $y \in \{x, Sx, SSx, SSSx, \dots\}$. We formalize this idea in the following definition:

DEFINITION. $\leq_{o\sigma\sigma}$ stands for

$$[\lambda x_\sigma \lambda y_\sigma \forall p_{o\sigma} \bullet [p_{o\sigma} x_\sigma \wedge \forall z_\sigma \bullet p_{o\sigma} z_\sigma \supset p_{o\sigma} \bullet Sz_\sigma] \supset p_{o\sigma} y_\sigma].$$

$\mathbf{A}_\sigma \leq \mathbf{B}_\sigma$ stands for $\leq_{o\sigma\sigma} \mathbf{A}_\sigma \mathbf{B}_\sigma$.

6200. $\vdash \dot{\forall} y_\sigma \bullet 0 \leq y_\sigma$

Proof:

- (.1) $\vdash \mathbb{N} y_\sigma = \bullet 0 \leq y_\sigma$ def of \mathbb{N} and \leq
 - (.2) $\vdash \dot{\forall} y_\sigma \bullet 0 \leq y_\sigma$.1
-

REMARK. Reflection on 6200.1 shows that we could have simply defined $\mathbb{N}_{o\sigma}$ as $\leq_{o\sigma\sigma} 0_\sigma$.

6201. $\vdash x_\sigma \leq x_\sigma$

Proof:

- (.1) $\vdash p_{o\sigma} x_\sigma \wedge \forall z_\sigma [pz \supset p \bullet Sz] \supset px$
 - (.2) $\vdash x_\sigma \leq x$.1, def of \leq
-

6202. $\vdash x_\sigma = y_\sigma \supset x_\sigma \leq y_\sigma$

Proof: By 6201. ■

6203. $\vdash x_\sigma \leq Sx_\sigma$

Proof:

- (.1) .1 $\vdash p_{o\sigma} x_\sigma \wedge \forall z_\sigma \bullet pz \supset p \bullet Sz$ hyp
 - (.2) .1 $\vdash p_{o\sigma} \bullet Sx_\sigma$.1
 - (.3) $\vdash x_\sigma \leq Sx$.2, def of \leq
-

6204. $\vdash x_\sigma \leq y_\sigma \wedge y_\sigma \leq z_\sigma \supset x_\sigma \leq z_\sigma$

Proof:

- (.1) .1 $\vdash x_\sigma \leq y_\sigma$ hyp
 - (.2) .2 $\vdash y_\sigma \leq z_\sigma$ hyp
 - (.3) .3 $\vdash p_{o\sigma} x_\sigma \wedge \forall w_\sigma \bullet pw \supset p \bullet Sw$ hyp
 - (.4) .1, .2, .3 $\vdash p_{o\sigma} y_\sigma$.1, .3, def of \leq
 - (.5) .1, .2, .3 $\vdash p_{o\sigma} z_\sigma$.2, .3, .4, def of \leq
 - (.6) .1, .2 $\vdash x_\sigma \leq z_\sigma$.5, def of \leq
-

6205. $\vdash \forall x_\sigma \cdot x_\sigma \leq y_\sigma \supset \neg x_\sigma = y_\sigma \vee \exists z_\sigma \cdot x_\sigma \leq z_\sigma \wedge Sz = y$

Proof: Let $P_{o\sigma}$ be $[\lambda w_\sigma \cdot x_\sigma = w \vee \exists z_\sigma \cdot x \leq z \wedge Sz = w]$.

(.1)	$\vdash x_\sigma \leq y_\sigma \supset \neg [Px \wedge \forall w_\sigma \cdot Pw \supset P \cdot Sw] \supset Py$	def of \leq
(.2)	$\vdash Px_\sigma$	5200, def of P
(.3)	$Pw_\sigma \vdash x_\sigma = w \vee \exists z_\sigma \cdot x \leq z \wedge Sz = w$	hyp, def of P
(.4)	$\mathbb{N}x_\sigma, x = w_\sigma \vdash \mathbb{N}x \wedge x \leq x \wedge Sx = Sw$	hyp, 6201
(.5)	$\mathbb{N}x_\sigma, x = w_\sigma \vdash \exists z_\sigma \cdot x \leq z \wedge Sz = Sw$	\exists Gen: .4
(.6)	$.6 \vdash \mathbb{N}z_\sigma \wedge x_\sigma \leq z \wedge Sz = w_\sigma$	hyp
(.7)	$.6 \vdash \mathbb{N}Sz_\sigma$.6, 6101
(.8)	$.6 \vdash x_\sigma \leq Sz_\sigma$.6, 6203, 6204
(.9)	$.6 \vdash SSz = Sw$.6
(.10)	$.6 \vdash \mathbb{N}Sz_\sigma \wedge x_\sigma \leq Sz \wedge SSz = Sw_\sigma$.7, .8, .9
(.11)	$.6 \vdash \exists z_\sigma \cdot x_\sigma \leq z \wedge Sz = Sw_\sigma$	\exists Gen: .10
(.12)	$\exists z_\sigma[.6] \vdash .11$.11
(.13)	$\mathbb{N}x_\sigma, Pw_\sigma \vdash .11$.3, .5, .12
(.14)	$\mathbb{N}x_\sigma \vdash \forall w_\sigma \cdot Pw \supset P \cdot Sw$.13, def of P
(.15)	$\mathbb{N}x_\sigma \vdash x \leq y_\sigma \supset Py$.1, .2, 14

■

6206. $\vdash \forall x_\sigma \forall y_\sigma \cdot x_\sigma \leq Sy_\sigma \equiv \neg x_\sigma = y_\sigma \vee x_\sigma = Sy_\sigma$

Proof:

(.1)	$.1 \vdash \mathbb{N}x_\sigma \wedge \mathbb{N}y_\sigma \wedge x_\sigma \leq Sy_\sigma$	hyp
(.2)	$.1 \vdash x_\sigma = Sy_\sigma \vee \exists z_\sigma \cdot x \leq z \wedge Sz = Sy$.1, 6205
(.3)	$.3 \vdash \mathbb{N}z_\sigma \wedge x_\sigma \leq z \wedge Sz = Sy$	hyp (choose z)
(.4)	$.1, .3 \vdash z_\sigma = y_\sigma$.1, .3, 6109
(.5)	$.1, .3 \vdash x_\sigma \leq y_\sigma$.3, .4
(.6)	$.1, \exists z_\sigma[.3] \vdash x_\sigma \leq y_\sigma$.5
(.7)	$.1 \vdash x_\sigma \leq y_\sigma \vee x = Sy$.2, .6
(.8)	$\vdash x_\sigma \leq y_\sigma \supset x \leq Sy$	6203, 6204
(.9)	$\vdash x_\sigma = Sy_\sigma \supset x \leq Sy$	6202
(.10)	$\vdash \mathbb{N}x_\sigma \wedge \mathbb{N}y_\sigma \supset \neg x \leq Sy \equiv \neg x = y \vee x = Sy$.7, .8, .9

■

6207⁰. $\vdash \forall x_\sigma \cdot x \leq 0 \equiv \neg x = 0$

6207¹. $\vdash \forall x_\sigma \cdot x \leq \bar{1} \equiv \neg x = 0 \vee x = \bar{1}$

6207^k. $\vdash \forall x_\sigma \cdot x \leq \bar{k} \equiv \neg x = 0 \vee x = \bar{1} \vee \dots \vee x = \bar{k}$

We prove these theorems by induction (in the meta-language) on k .

- (.1) $\vdash \mathbb{N}x_\sigma \supset \blacksquare x \leq 0 \supset \blacksquare x = 0 \vee \dot{\exists} z_\sigma \blacksquare x \leq z \wedge Sz = 0$ 6205
(.2) $\vdash \sim \dot{\exists} z_\sigma \blacksquare x_\sigma \leq z \wedge Sz = 0$ 6106
(.3) $\vdash \dot{\forall} x_\sigma \blacksquare x \leq 0 \equiv \blacksquare x = 0$.1, .2, 6202

This proves 6207^0 . Next we assume we have proved 6207^k (where $k \geq 0$), and prove 6207^{k+1} .

- (.4) $\mathbb{N}x_\sigma \vdash x \leq \bar{k} \equiv \blacksquare x = 0 \vee \dots \vee x = \bar{k}$ by inductive hypothesis
(.5) $\mathbb{N}x_\sigma \vdash x \leq S\bar{k} \equiv \blacksquare x \leq \bar{k} \vee x = S\bar{k}$ 6206, 6110
(.6) $\mathbb{N}x_\sigma \vdash x \leq S\bar{k} \equiv \blacksquare x = 0 \vee \dots \vee x = \bar{k} \vee x = S\bar{k}$.4, .5

This gives the desired theorem, since $\overline{k+1}$ is $S\bar{k}$. ■

6208 Theorem. $\leq_{o\sigma\sigma}$ represents the relation \leq between natural numbers.

Proof: $\vdash \bar{n} \leq \bar{k} = \blacksquare \bar{n} = 0 \vee \dots \vee \bar{n} = \bar{k}$, and $\mathbf{Q}_{o\sigma\sigma}$ represents equality (6112). ■

6209^k. $\vdash \dot{\forall} y_\sigma [y_\sigma \leq \bar{k} \supset p_{o\sigma} y_\sigma] = \blacksquare p_{o\sigma} 0 \wedge \dots \wedge p_{o\sigma} \bar{k}$

Proof: Let $j \leq k$ in each line below where j occurs.

- (.1^j) $\vdash \dot{\forall} y_\sigma [y \leq \bar{k} \supset p_{o\sigma} y] \supset \blacksquare j \leq \bar{k} \supset p_j$ 6110
(.2^j) $\vdash \bar{j} \leq \bar{k}$ 6208
(.3) $\vdash \dot{\forall} y_\sigma [y \leq \bar{k} \supset p_{o\sigma} y] \supset \blacksquare p_0 \wedge \dots \wedge p_{\bar{k}}$.1⁰ – .1^k, .2⁰ – .2^k
(.4) .4 $\vdash p_{o\sigma} 0 \wedge \dots \wedge p_{\bar{k}} \wedge \mathbb{N}y_\sigma \wedge y \leq \bar{k}$ hyp
(.5) .4 $\vdash y_\sigma = 0 \vee \dots \vee y = \bar{k}$.4, 6207^k
(.6^j) $\vdash y_\sigma = \bar{j} \supset \blacksquare p_{o\sigma} y = p_j$ Axiom 2
(.7) .4 $\vdash p_{o\sigma} y_\sigma$.4, .5, .6⁰ – .6^k
(.8) $\vdash p_{o\sigma} 0 \wedge \dots \wedge p_{\bar{k}} \supset \dot{\forall} y_\sigma \blacksquare y \leq \bar{k} \supset p_y$.7

The theorem follows from .3 and .8. ■

6210^k. $\vdash \dot{\exists} y_\sigma [y_\sigma \leq \bar{k} \wedge p_{o\sigma} y_\sigma] = \blacksquare p_{o\sigma} 0 \vee \dots \vee p_{o\sigma} \bar{k}$

Proof:

- (.1) $\vdash \dot{\forall} y_\sigma [y \leq \bar{k} \supset \sim p_{o\sigma} y] = \blacksquare \sim p_0 \wedge \dots \wedge \sim p_{\bar{k}}$ Sub: 6209^k
(.2) $\vdash \sim \dot{\forall} y_\sigma [y \leq \bar{k} \supset \sim p_{o\sigma} y] = \dot{\exists} y_\sigma \blacksquare y \leq \bar{k} \wedge p_y$ 6000, defns of $\dot{\forall}$ and $\dot{\exists}$

The theorem follows from .1 and .2 by Rule P. ■

6211. $\vdash \dot{\forall} y_\sigma \dot{\forall} x_\sigma \leq y_\sigma \supset Sx_\sigma \leq Sy_\sigma$

Proof by induction on y:

- (.1) .1 $\vdash \mathbb{N} x_\sigma \wedge x \leq 0$ hyp
- (.2) .1 $\vdash x_\sigma = 0$.1, 6207⁰
- (.3) .1 $\vdash Sx_\sigma = S0$.2
- (.4) .1 $\vdash Sx_\sigma \leq S0$.3, 6202
- (.5) $\vdash \dot{\forall} x_\sigma \bullet x \leq 0 \supset Sx \leq S0$.4
- (.6) .6 $\vdash \mathbb{N} y_\sigma \wedge \dot{\forall} x_\sigma \bullet x \leq y_\sigma \supset Sx \leq Sy$ (inductive) hyp
- (.7) .7 $\vdash \mathbb{N} x_\sigma \wedge x \leq Sy$ hyp
- (.8) .6, .7 $\vdash x_\sigma \leq y_\sigma \vee x = Sy$.6, .7, 6206
- (.9) .6, .7, $x_\sigma \leq y_\sigma \vdash Sx \leq Sy$ hyp, .6, .7
- (.10) .6, .7, $x_\sigma \leq y_\sigma \vdash Sx \leq SSy$.9, 6203, 6204
- (.11) .6, .7, $x_\sigma = Sy_\sigma \vdash Sx = SSy$ hyp
- (.12) .6, .7, $x_\sigma = Sy_\sigma \vdash Sx \leq SSy$.11, 6202
- (.13) .6, .7 $\vdash Sx_\sigma \leq SSy_\sigma$.8, .10, .12
- (.14) .6 $\vdash \dot{\forall} x_\sigma \bullet x \leq Sy_\sigma \supset Sx \leq SSy$.13
- (.15) $\vdash \dot{\forall} y_\sigma \dot{\forall} x_\sigma \bullet x \leq y \supset Sx \leq Sy$ 6102, .5, .14

■

6212. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \bullet x \leq y \vee y \leq x$

Proof by induction on x:

- (.1) $\vdash \dot{\forall} y_\sigma \bullet 0 \leq y \vee y \leq 0$ 6200
- (.2) .2 $\vdash \mathbb{N} x_\sigma \wedge \dot{\forall} y_\sigma \bullet x \leq y \vee y \leq x$ (inductive) hyp
- (.3) .2, $\mathbb{N} y_\sigma \vdash x_\sigma \leq y \vee y \leq x$.2, hyp
- (.4) $\vdash y_\sigma \leq x_\sigma \supset y \leq Sx$ 6203, 6204
- (.5) .2 $\vdash x_\sigma \leq y_\sigma \supset \bullet x = y \vee \dot{\exists} z_\sigma \bullet x \leq z \wedge Sz = y$.2, 6205
- (.6) $\vdash x_\sigma = y_\sigma \supset y \leq Sx$ 5302, 6202, .4
- (.7) .7 $\vdash \mathbb{N} z_\sigma \wedge x_\sigma \leq z \wedge Sz = y$ hyp (choose z)
- (.8) .2, .7 $\vdash Sx_\sigma \leq Sz_\sigma$.2, .7, 6211
- (.9) .2, .7 $\vdash Sx_\sigma \leq y_\sigma$.7, .8
- (.10) .2, $\dot{\exists} z_\sigma[.7] \vdash Sx_\sigma \leq y_\sigma$.9
- (.11) .2, $\mathbb{N} y_\sigma \vdash Sx_\sigma \leq y \vee y \leq Sx$.3, .4, .5, .6, .10
- (.12) $\vdash \dot{\forall} x_\sigma \bullet \dot{\forall} y_\sigma[x \leq y \vee y \leq x] \supset \dot{\forall} y_\sigma \bullet Sx \leq y \vee y \leq Sx$.11

The theorem follows from .1 and .12 by the Induction Theorem. ■

6213. $\vdash \dot{\forall} y_\sigma \dot{\forall} x_\sigma \bullet x \leq y \wedge y \leq x \supset x = y$

Proof by induction on y:

- (.1) $\vdash \dot{\forall} x_\sigma \bullet x \leq 0 \wedge 0 \leq x \supset x = 0$ 6207⁰
- (.2) .2 $\vdash \dot{\mathbb{N}} y_\sigma \wedge \dot{\forall} x_\sigma \bullet x \leq y \wedge y \leq x \supset x = y$ (inductive) hyp
- (.3) .3 $\vdash \dot{\mathbb{N}} x_\sigma \wedge x \leq S y_\sigma \wedge S y_\sigma \leq x$ hyp
- (.4) .2, .3 $\vdash x_\sigma \leq y_\sigma \vee x = S y$.3, 6206
- (.5) .5 $\vdash x_\sigma \leq y_\sigma$ hyp
- (.6) .2, .3, .5 $\vdash y_\sigma \leq x_\sigma$.3, 6203, 6204
- (.7) .2, .3, .5 $\vdash x_\sigma = y_\sigma$.2, .3, .5, .6

Intuitively, .3 and .7 are contradictory, so under the assumptions .2 and .3, we have found that .5 is impossible, so $x = S y$ by .4. Actually, it proves easier to derive $x = S y$ than to derive an actual contradiction.

- (.8) .2, .3, .5 $\vdash y_\sigma \leq S x_\sigma \wedge S x \leq y$ R': .7, .3
- (.9) .2, .3, .5 $\vdash \dot{\mathbb{N}} \bullet S x_\sigma$.3, 6101
- (.10) .2, .3, .5 $\vdash S x_\sigma = y_\sigma$.2, .8, .9
- (.11) .2, .3, .5 $\vdash S y_\sigma = x_\sigma$ R': .7, .10
- (.12) .2, .3 $\vdash x_\sigma = S y_\sigma$.4, .11, 5302
- (.13) $\vdash \dot{\forall} y_\sigma \bullet \dot{\forall} x_\sigma [x \leq y \wedge y \leq x \supset x = y]$
 $\supset \dot{\forall} x_\sigma \bullet x \leq S y \wedge S y \leq x \supset x = S y$.12

The theorem follows from .1 and .13 by the Induction Theorem. ■

6214. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \bullet \sim x \leq y \supset S y \leq x$

Proof:

- (.1) .1 $\vdash \dot{\mathbb{N}} x_\sigma \wedge \dot{\mathbb{N}} y_\sigma \wedge \sim x \leq y$ hyp
- (.2) .1 $\vdash x_\sigma \leq S y_\sigma \vee S y \leq x$.1, 6101, 6212
- (.3) .1 $\vdash x_\sigma \leq S y_\sigma \equiv \bullet x \leq y \vee x = S y$.1, 6206
- (.4) .1 $\vdash x_\sigma = S y_\sigma \supset S y \leq x$ 5302, 6202
- (.5) .1 $\vdash S y_\sigma \leq x_\sigma$.1, .2, .3, .4

■

We may now supply more details of the discussion about axioms of infinity begun in §60. Let $Infin^\alpha$ be the wff

$$\exists r_{o\alpha\alpha} \forall x_\alpha \forall y_\alpha \forall z_\alpha [\exists w_\alpha rxw \wedge \sim rxx \wedge \neg rxy \supset \neg ryx \supset rxz].$$

Note that $Infin^i$ is Axiom 6.

6215. $\vdash Infin^\sigma$

Proof: Let $R_{\sigma\sigma\sigma}$ be the wff $[\lambda n_\sigma \lambda m_\sigma \cdot \mathbb{N}m \wedge \sim \mathbb{N}n \vee Sn \leq m]$.

Note that $Sn \leq m$ means that $n < m$.

(.1)	$\vdash \sim Nx_\sigma \supset Rx0$	6100, def of R
(.2)	$\vdash \sim Nx_\sigma \supset \exists w_\sigma Rxw$.1
(.3)	$\vdash Nx_\sigma \supset Rx \cdot Sx$	6101, 6201, def of R
(.4)	$\vdash Nx_\sigma \supset \exists w_\sigma Rxw$.3
(.5)	$\vdash \exists w_\sigma Rx_\sigma w$.2, .4
(.6)	$.6 \vdash Rx_\sigma x$	hyp
(.7)	$.6 \vdash Nx_\sigma \wedge Sx \leq x$.6, def of R
(.8)	$.6 \vdash x_\sigma = Sx$.7, 6203, 6101, 6213
(.9)	$.6 \vdash F_o$.7, .8, 6111
(.10)	$\vdash \sim Rx_\sigma x$.9
(.11)	$.11 \vdash Rx_\sigma y_\sigma \wedge Ryx_\sigma$	hyp
(.12)	$.11 \vdash Ny_\sigma \wedge Nz_\sigma$.11, def of R
(.13)	$.11 \vdash \sim Nx_\sigma \supset Rxz_\sigma$.12, def of R
(.14)	$.11, .14 \vdash Nx_\sigma$	hyp
(.15)	$.11, .14 \vdash Sx_\sigma \leq y_\sigma \wedge Sy \leq z_\sigma$.11, .14, def of R
(.16)	$.11, .14 \vdash Sx_\sigma \leq z_\sigma$.15, 6203, 6204
(.17)	$.11 \vdash Nx_\sigma \supset Rx_\sigma z_\sigma$.12, .16, def of R
(.18)	$\vdash Rx_\sigma y_\sigma \wedge Ryx_\sigma \supset Rxz$.13, .17
(.19)	$\vdash Infin^\sigma$.5, .10, .18

■

Note that we have now shown that $\vdash_{Q_0} Infin^i \supset Infin^\sigma$. It is clear that $\vdash_{Q_0} Infin^\alpha \supset Infin^{(o(o\alpha))}$. Also, we have shown that without using the definitions of σ , 0, or S , one can prove $\vdash_{Q_0} [6106] \wedge [6109] \supset Infin^\sigma$. (See the discussion following the proof of 6109. For the sake of economy we have used 6105 in several places where we could have used 6106.) The reader will now find it easy to fill in the details of the argument in §60 that $\vdash_{Q_0} Inf^{II} \supset$ Axiom 6.

EXERCISES

X6200. Note that we have used the definition of \leq only in the proofs of theorems 6200, 6201, 6203, 6204, and 6205. Of course, there are many theorems about \leq which remain to be proved. The question naturally arises whether the definition of \leq must be used (directly) again to derive some of these, or whether 6200–6205 characterize \leq on the natural numbers. (In the latter case, of course, first-order formulations of these theorems could be taken as axioms in a first-order theory of natural numbers in which \leq is a primitive predicate constant.) Investigate this question by considering whether the following wff is a theorem:

$$\forall r_{o\sigma} \cdot \dot{\forall} x_\sigma \dot{\forall} y_\sigma \dot{\forall} z_\sigma [r0y \wedge rxx \wedge rx[Sx] \wedge [rxy \wedge ryz \supset rxz] \\ \wedge \neg rxy \supset \neg x = y \vee \exists w_\sigma \cdot rxw \wedge Sw = y] \supset \dot{\forall} x_\sigma \dot{\forall} y_\sigma \cdot rxy \equiv x \leq y.$$

X6201. Show that $\vdash_{Q_0} Infin^\alpha \supset Infin^{(\alpha)}$.

X6202. Let \max be the numerical function such that

$$\begin{aligned} \max(x, y) &= x \text{ if } x \geq y; \\ &= y \text{ otherwise.} \end{aligned}$$

Show that \max is a representable function.

§63. Minimization

The theorems of §62 show that \leq orders the set \mathbb{N} of natural numbers. We shall now prove that this ordering is a well-ordering, i.e., that every nonempty subset p of \mathbb{N} has a least member μp . Indeed, we shall define a mapping $\mu_{\sigma(o\sigma)}$ which maps each set $p_{o\sigma}$ to the least natural number in $p_{o\sigma}$ (if there is one). Thus, $\mu_{\sigma(o\sigma)}$ will be a choice function for \mathbb{N} .

DEFINITIONS.

$\mu_{\sigma(o\sigma)}$ stands for $[\lambda p_{o\sigma} \cdot x_\sigma \cdot \mathbb{N}_{o\sigma} x_\sigma \wedge p_{o\sigma} x_\sigma \wedge \dot{\forall} y_\sigma \cdot p_{o\sigma} y_\sigma \supset x_\sigma \leq y_\sigma]$.
 $[\mu x_\sigma A]$ stands for $\mu_{\sigma(o\sigma)}[\lambda x_\sigma A]$.

$[\mu x_\sigma A]$ denotes “the least natural number x_σ such that A ” (when one exists). For this reason, μ is sometimes called “the least-number operator”.

6300. $\vdash \mathbb{N}_{o\sigma} t_\sigma \wedge p_{o\sigma} t_\sigma \supset \bullet \mathbb{N}_{o\sigma} [\mu_{\sigma(o\sigma)} p_{o\sigma}] \wedge p_{o\sigma} [\mu_{\sigma(o\sigma)} p_{o\sigma}] \wedge \mu_{\sigma(o\sigma)} p_{o\sigma} \leq t_\sigma$

Proof: Let $G_{o\sigma}$ be $[\lambda x_\sigma \bullet \mathbb{N}x \wedge px \wedge \dot{\forall}y_\sigma \bullet py \supset x \leq y]$. Note that $p_{o\sigma}$ is the sole free variable in $G_{o\sigma}$, which denotes the set of smallest natural numbers in the set denoted by $p_{o\sigma}$.

$$(1) \vdash \mu_{\sigma(o\sigma)} p_{o\sigma} = \iota_{\sigma(o\sigma)} G_{o\sigma} \quad \text{defs of } \mu \text{ and } G$$

We wish to prove $\exists_1 x_\sigma G_{o\sigma} x_\sigma$ from the hypothesis $\mathbb{N}_{o\sigma} t_\sigma \wedge p_{o\sigma} t_\sigma$. Intuitively, it is clear that if t_σ is a natural number in $p_{o\sigma}$, then there is a least natural number in $p_{o\sigma}$ and it will be in $G_{o\sigma}$. How do we “find” this number? We can imagine examining natural numbers $\leq t_\sigma$ in order until we first come upon one in $p_{o\sigma}$. This suggests that some kind of an inductive process may be useful. Of course, a proof by induction will work only if we seek to prove a statement which is actually true for all natural numbers. Therefore, let us prove $\dot{\forall}z_\sigma \bullet \dot{\exists}x_\sigma [p_{o\sigma} x_\sigma \wedge x_\sigma \leq z_\sigma] \supset \exists x_\sigma G_{o\sigma} x_\sigma$ by induction on z .

- | | | |
|------|--|-----------------------|
| (.2) | $.2 \vdash \mathbb{N}x_\sigma \wedge p_{o\sigma} x \wedge x \leq 0$ | hyp |
| (.3) | $.2 \vdash x_\sigma = 0$ | .2, 6207 ⁰ |
| (.4) | $.2 \vdash \dot{\forall}y_\sigma \bullet p_{o\sigma} y \supset x_\sigma \leq y$ | .3, 6200 |
| (.5) | $.2 \vdash Gx_\sigma$ | .2, .4, def of G |
| (.6) | $\vdash \dot{\exists}x_\sigma [p_{o\sigma} x \wedge x \leq 0] \supset \exists x_\sigma Gx$ | .5 |

Next we treat the induction step.

- | | | |
|------|---|-----------------|
| (.7) | $.7 \vdash \mathbb{N}z_\sigma \wedge \dot{\exists}x_\sigma [p_{o\sigma} x \wedge x \leq z] \supset \exists x_\sigma Gx$ | (inductive) hyp |
| (.8) | $.8 \vdash \mathbb{N}x_\sigma \wedge p_{o\sigma} x \wedge x \leq Sz_\sigma$ | hyp |

From .7 and .8 we must prove $\exists x_\sigma Gx$. From .7 it can be seen that we really need be concerned only with the case in line .9.

- | | | |
|-------|--|---------------------|
| (.9) | $.7, .8, .9 \vdash \dot{\exists}x_\sigma \bullet p_{o\sigma} x \wedge x \leq z_\sigma$ | hyp |
| (.10) | $.7, .8, .9 \vdash \dot{\forall}y_\sigma \bullet p_{o\sigma} y \supset \sim y \leq z_\sigma$ | .9 |
| (.11) | $.7, .8, .9, .11 \vdash \mathbb{N}y_\sigma \wedge p_{o\sigma} y_\sigma$ | hyp |
| (.12) | $.7, .8, .9, .11 \vdash \sim y \leq z_\sigma$ | .10, .11 |
| (.13) | $.7, .8, .9, .11 \vdash Sz_\sigma \leq y_\sigma$ | .7, .11, .12, 6214 |
| (.14) | $.7, .8, .9, .11 \vdash x_\sigma \leq y_\sigma$ | .8, .13, 6204 |
| (.15) | $.7, .8, .9 \vdash \dot{\forall}y_\sigma \bullet p_{o\sigma} y \supset x \leq y$ | .14 |
| (.16) | $.7, .8, .9 \vdash Gx_\sigma$ | .8, .15, def of G |
| (.17) | $.7, .8, .9 \vdash \exists x_\sigma Gx$ | .16 |
| (.18) | $.7, .8, \vdash \exists x_\sigma Gx$ | .7, .17 |

- (.19) .7, $\exists x_\sigma[.8] \vdash \exists x_\sigma Gx$.18
(.20) $\vdash \forall z_\sigma \bullet \dot{\exists} x_\sigma[p_{o\sigma}x \wedge x \leq z] \supset \exists x_\sigma Gx \supset \bullet \dot{\exists} x_\sigma[p_{o\sigma}x \wedge x \leq Sz] \supset \exists x_\sigma Gx$.19

This completes the induction step.

- (.21) $\vdash \dot{\forall} z_\sigma \bullet \dot{\exists} x_\sigma[p_{o\sigma}x \wedge x \leq z] \supset \exists x_\sigma Gx$ 6102, .6, .20
(.22) .22 $\vdash \mathbb{N}t_\sigma \wedge p_{o\sigma}t$ hyp
(.23) .22 $\vdash t_\sigma \leq t$ 6201
(.24) .22 $\vdash \dot{\exists} x_\sigma \bullet p_{o\sigma}x \wedge x \leq t_\sigma$.22, .23
(.25) .22 $\vdash \exists x_\sigma Gx$.21, .22, .24

Next we prove the uniqueness of membership in $G_{o\sigma}$.

- (.26) .26 $\vdash Gx_\sigma \wedge Gz_\sigma$ hyp
(.27) .26 $\vdash \mathbb{N}x_\sigma \wedge p_{o\sigma}x \wedge \mathbb{N}z_\sigma \wedge pz$.26
(.28) .26 $\vdash \dot{\forall} y_\sigma \bullet p_{o\sigma}y \supset x_\sigma \leq y$.26
(.29) .26 $\vdash \dot{\forall} y_\sigma \bullet p_{o\sigma}y \supset z_\sigma \leq y$.26
(.30) .26 $\vdash x_\sigma \leq z_\sigma \wedge z \leq x$.27, .28, .29
(.31) .26 $\vdash x_\sigma = z_\sigma$ 6213, .30, .27
(.32) $\vdash \forall x_\sigma \forall z_\sigma \bullet Gx \wedge Gz \supset x = z$.31
(.33) .22 $\vdash \exists_1 x_\sigma Gx$ 5307, .25, .32
(.34) .22 $\vdash G_{o\sigma} \bullet \iota_{\sigma(o\sigma)}G_{o\sigma}$ 5311, .33
(.35) .22 $\vdash G_{o\sigma} \bullet \mu_{\sigma(o\sigma)}p_{o\sigma}$.1, .34
(.36) .22 $\vdash \mathbb{N}\mu p_{o\sigma} \wedge p\mu p \wedge \dot{\forall} y_\sigma \bullet py \supset \mu p \leq y$.35, def of G
(.37) .22 $\vdash \mu p_{o\sigma} \leq t_\sigma$.22, .36

The theorem follows from .36 and .37. ■

6301. $\vdash \mathbb{N}_{o\sigma}t_\sigma \wedge p_{o\sigma}t_\sigma \wedge \dot{\forall} y_\sigma[y_\sigma \leq t_\sigma \supset \bullet p_{o\sigma}y_\sigma \supset y_\sigma = t_\sigma] \supset \mu_{\sigma(o\sigma)}p_{o\sigma} = t_\sigma$

Proof:

- (.1) .1 $\vdash \mathbb{N}t_\sigma \wedge p_{o\sigma}t \wedge \dot{\forall} y_\sigma \bullet y \leq t \supset \bullet py \supset y = t$ hyp
(.2) .1 $\vdash \mathbb{N}\mu p_{o\sigma} \wedge p\mu p \wedge \mu p \leq t_\sigma$ 6300, .1
(.3) .1 $\vdash \mu p_{o\sigma} = t_\sigma$ Rule P, VI: .1, .2 ■

We now need to introduce some terminology. We shall refer to functions from natural numbers to natural numbers as *numerical functions*. For the

sake of uniformity of terminology, we shall sometimes refer to a natural number as a numerical function of zero arguments. Relations between natural numbers shall be referred to as *numerical relations*.

It will be convenient to use λ -notation in our meta-language to define numerical functions and relations. Thus, if G is a numerical expression (such as $2 \times n_1 \times (n_2 + n_3)$) in our meta-language, possibly containing numerical parameters n_1, \dots, n_k , we let $((\lambda n_1, \dots, n_k)G)$ denote that k -ary numerical function f such that $f(n_1, \dots, n_k) = G$ for all natural numbers n_1, \dots, n_k . Similarly, if B is a statement (such as $n_1 + n_2 = n_3$) in our meta-language about natural numbers n_1, \dots, n_k , we let $((\lambda n_1, \dots, n_k)B)$ denote that k -ary numerical relation r such that $r(n_1, \dots, n_k)$ iff B , for all natural numbers n_1, \dots, n_k . (Of course, in other contexts one might use λ -notation to denote functions and relations defined over domains other than the set of natural numbers.)

We shall also use the least-number operator μ in our meta-language. If B is a statement about a number n , then $(\mu n)B$ shall denote the least natural number n for which B is true, if one exists; if none exists, then $(\mu n)B$ will be undefined.

As an example of these notations, $((\lambda n)(\mu m)n \leq m^2)$ denotes that numerical function f whose value at n is the least natural number $\geq \sqrt{n}$; in particular if $n = k^2$, $f(n) = k = \sqrt{n}$.

In principle, we could also introduce the description operator τ into our meta-language. However, this is unnecessary as long as we confine our attention to natural numbers, since if there is a unique natural number n such that B , then “that n such that B ” is “the least n such that B ”.

A $(k+1)$ -ary numerical function g is said to be *regular* iff for all natural numbers n_1, \dots, n_k there is a natural number m such that $g(n_1, \dots, n_k, m) = 1$. A $(k+1)$ -ary numerical relation r is said to be *regular* iff its characteristic function χ_r is regular.

If g is a regular $(k+1)$ -ary numerical function and

$$h = ((\lambda n_1, \dots, n_k)(\mu m)g(n_1, \dots, n_k, m)) = 1,$$

we say that h is obtained from g by *minimization*. Similarly, if r is a regular $(k+1)$ -ary numerical relation and

$$h = ((\lambda n_1, \dots, n_k)(\mu m)r(n_1, \dots, n_k, m)),$$

we say that h is obtained from r by *minimization*. Note that the regularity of g (or of r) assures that h is defined everywhere. Also note that h is obtained from a relation r by minimization iff h is obtained from χ_r by minimization.

6302 Theorem. If r is a representable $(k+1)$ -ary regular numerical relation, then $((\lambda n_1, \dots, n_k)(\mu m)r(n_1, \dots, n_k, m))$ is a representable function.

Proof: Let $h = ((\lambda n_1, \dots, n_k)(\mu m)r(n_1, \dots, n_k, m))$. Suppose $R_{o\sigma\sigma\dots\sigma}$ represents r , and let $H_{\sigma\sigma\dots\sigma}$ be $[\lambda x_\sigma^1 \dots \lambda x_\sigma^k \mu_{\sigma(\sigma\sigma)} \bullet R_{o\sigma\sigma\dots\sigma} x_\sigma^1 \dots x_\sigma^k]$. We show that H represents h . Suppose $h(n_1, \dots, n_k) = m$. Thus

- (.1) $\vdash R\bar{n_1} \dots \bar{n_k}\bar{m} = T_o$ and
 - (.2) $\vdash R\bar{n_1} \dots \bar{n_k}\bar{j} = F_o$ for each $j < m$.
 - (.3) $\vdash \bar{N}\bar{m} \wedge R\bar{n_1} \dots \bar{n_k}\bar{m} \wedge \forall y_\sigma [y_\sigma \leq \bar{m} \supset \bullet R\bar{n_1} \dots \bar{n_k}y \supset \bullet y = \bar{m}] \supset$
- 6301
 $\bullet \mu R\bar{n_1} \dots \bar{n_k} = \bar{m}$

Let $P_{o\sigma}$ be $[\lambda y_\sigma \bullet R\bar{n_1} \dots \bar{n_k}y_\sigma \supset y_\sigma = \bar{m}]$.

- (.4) $\vdash \forall y_\sigma [y \leq \bar{m} \supset P_{o\sigma}y_\sigma] = \bullet P_{o\sigma}0 \wedge \dots \wedge P_{o\sigma}\bar{m}$ 6209^m
 - (.5) $\vdash P_{o\sigma}\bar{j}$ for each $j < m$.2, def of P
 - (.6) $\vdash P_{o\sigma}\bar{m}$ 5200, def of P
 - (.7) $\vdash \mu R\bar{n_1} \dots \bar{n_k} = \bar{m}$.3, 6110, .1, .4, .5, .6, def of P
 - (.8) $\vdash H_{\sigma\sigma\dots\sigma}\bar{n_1} \dots \bar{n_k} = \bar{m}$ def of H , .7
-

6303 Theorem. If g is a representable $(k+1)$ -ary regular numerical function, and h is obtained from g by minimization, then h is representable.

Proof: Let the relation r be $(\lambda n_1, \dots, n_k, m)(g(n_1, \dots, n_k, m) = 1)$. Clearly r is regular, so by 6302 it suffices to show that r is representable. Let $G_{\sigma\sigma\sigma\dots\sigma}$ represent g , and let $R_{o\sigma\sigma\dots\sigma}$ be $[\lambda x_\sigma^1 \dots \lambda x_\sigma^k \lambda y_\sigma \bullet Gx^1 \dots x^ky = \bar{1}]$. We show that R represents r . Let n_1, \dots, n_k, m be any natural numbers.

- (.1) $\vdash R\bar{n_1} \dots \bar{n_k}\bar{m} \equiv \bullet G\bar{n_1} \dots \bar{n_k}\bar{m} = \bar{1}$ def of R
- (.2) $\vdash G\bar{n_1} \dots \bar{n_k}\bar{m} = g(n_1, \dots, n_k, m)$ since G represents g

Suppose $r(n_1, \dots, n_k, m)$. Then $g(n_1, \dots, n_k, m) = 1$.

- (.3) $\vdash R\bar{n_1} \dots \bar{n_k}\bar{m} = T_o$.1, .2

Suppose not $r(n_1, \dots, n_k, m)$. Then $g(n_1, \dots, n_k, m) \neq 1$.

- (.4) $\vdash \overline{g(n_1, \dots, n_k, m)} \neq \bar{1}$ 6112
 - (.5) $\vdash R\bar{n_1} \dots \bar{n_k}\bar{m} = F_o$.1, .2, .4
-

§64. Recursive Functions

DEFINITION. Let g be a k -ary numerical function (where $k \geq 0$), and h be a $(k+2)$ -ary numerical function. A $(k+1)$ -ary function f is said to be defined by *primitive recursion* from g and h iff the following equations hold for all natural numbers m_1, \dots, m_k, n :

$$f(m_1, \dots, m_k, 0) = g(m_1, \dots, m_k)$$

$$f(m_1, \dots, m_k, Sn) = h(m_1, \dots, m_k, n, f(m_1, \dots, m_k, n))$$

EXAMPLE: $m + 0 = m$ and $m + Sn = S(m + n)$ for all natural numbers m and n , where S is the successor function. Thus, if we let $g(m) = m$ and $h(m, n, p) = Sp$ for all natural numbers m , n , and p , and define

$$f(m, 0) = g(m)$$

$$f(m, Sn) = h(m, n, f(m, n)),$$

then $f(m, n) = m + n$ for all natural numbers m and n , so addition is defined by primitive recursion from g and h .

It can be seen that for given numerical functions g and h , there is a unique numerical function f satisfying these equations. Indeed, if f and f' both satisfy the equations, then it is easy to prove by induction on n that $f(m_1, \dots, m_k, n) = f'(m_1, \dots, m_k, n)$ for all natural numbers m_1, \dots, m_k , and n , so $f = f'$. The existence of a function f satisfying these equations may seem more obvious at first glance than is really justified. (For an excellent discussion of this matter, see [Henkin, 1960].) In any case, we shall give a proof within Q_0^∞ of the existence of such a function. Indeed, since f is uniquely determined by h and g , we can define a recursion operator R such that $f = Rhg$. Since m_1, \dots, m_k can be regarded as parameters, it suffices to define R for the case where $k = 0$ (so g is simply a natural number).

Before giving the formal definition of R , let us provide an informal explanation of the method to be used. R yields f as a function of h and g , so it suffices to define f in terms of h and g . To do this it suffices to define the set C of ordered pairs $\langle n, m \rangle$ such that $f(n) = m$. Since $f(0) = g$, the pair $\langle 0, g \rangle$ is in C . Also, whenever $\langle n, m \rangle \in C$, we know that $f(n) = m$, so $f(Sn) = h(n, f(n)) = h(n, m)$, so $\langle Sn, h(n, m) \rangle \in C$. Thus, C can be defined inductively by the conditions that $\langle 0, g \rangle \in C$ and

that $\forall n \forall m [\langle n, m \rangle \in C \Rightarrow \langle S_n, h(n, m) \rangle \in C]$. This leads to the following definition:

DEFINITION. $R_{\sigma\sigma\sigma(\sigma\sigma\sigma)}$ stands for $[\lambda h_{\sigma\sigma\sigma} \lambda g_\sigma \lambda n_\sigma \lambda m_\sigma \forall w_{\sigma\sigma\sigma} \bullet [w_{\sigma\sigma\sigma} 0_\sigma g_\sigma \wedge \forall x_\sigma \forall y_\sigma \bullet w_{\sigma\sigma\sigma} x_\sigma y_\sigma \supset w_{\sigma\sigma\sigma} [S_{\sigma\sigma} x_\sigma] h_{\sigma\sigma\sigma} x_\sigma y_\sigma] \supset w_{\sigma\sigma\sigma} n_\sigma m_\sigma]$.

6400. $\vdash [R_{\sigma\sigma\sigma(\sigma\sigma\sigma)} h_{\sigma\sigma\sigma} g_\sigma] 0_\sigma = g_\sigma \wedge \forall n_\sigma \bullet [Rh_{\sigma\sigma\sigma} g_\sigma] [S_{\sigma\sigma} n] = h_{\sigma\sigma\sigma} n_\sigma \bullet [Rh_{\sigma\sigma\sigma} g_\sigma] n_\sigma$

Proof: Let $B_{o(o\sigma\sigma)}$ be $[\lambda w_{o\sigma\sigma} \bullet \forall x_\sigma \forall y_\sigma \bullet w_{o\sigma\sigma} x_\sigma y_\sigma \supset w_{o\sigma\sigma} [S_{\sigma\sigma} x_\sigma] h_{\sigma\sigma\sigma} x_\sigma y_\sigma]$ and let $C_{o\sigma\sigma}$ be $[\lambda n_\sigma \lambda m_\sigma \bullet \forall w_{o\sigma\sigma} \bullet w_{o\sigma\sigma} 0_\sigma g_\sigma \wedge Bw_{o\sigma\sigma} \supset w_{o\sigma\sigma} n_\sigma m_\sigma]$. Note that the free variables of $C_{o\sigma\sigma}$ are g_σ and $h_{\sigma\sigma\sigma}$. Also note that if $W_{o\sigma\sigma}$ is any wff such that $W 0 g_\sigma \wedge BW$ is derivable from certain hypotheses, then $Cxy \supset Wxy$ is also derivable from those hypotheses. We shall use this fact twice below.

$$(1) \vdash R_{\sigma\sigma\sigma(\sigma\sigma\sigma)} h_{\sigma\sigma\sigma} g_\sigma n_\sigma = \iota_{\sigma(\sigma\sigma)} [C_{o\sigma\sigma} n_\sigma] \quad \lambda, \text{ defns of } R, C$$

Intuitively, $Cn_\sigma m_\sigma$ means that m_σ is a possible value for $Rhgn$. We first prove that $\forall n_\sigma \exists m_\sigma [Cnm \wedge \forall k_\sigma \bullet Cnk \supset m = k]$ by induction on n .

$$(2) \vdash \exists m_\sigma [C0m \wedge \forall k_\sigma \bullet C0k \supset m = k] \wedge \begin{aligned} & \forall n_\sigma [\exists m_\sigma [Cnm \wedge \forall k_\sigma \bullet Cnk \supset m = k]] \\ & \supset \exists m_\sigma [C[S_n]m \wedge \forall k_\sigma \bullet C[S_n]k \supset m = k]] \\ & \supset \forall n_\sigma \exists m_\sigma \bullet Cnm \wedge \forall k_\sigma \bullet Cnk \supset m = k \end{aligned} \quad 6102$$

Essentially we are seeking to prove that $\forall n \exists m Cnm$. First we establish theorems that will enable us to prove $\exists m Cnm$.

$$\begin{aligned} (3) \vdash C0g_\sigma & \quad \text{def of } C \\ (4) Cn_\sigma m_\sigma, w_{o\sigma\sigma} 0g_\sigma \wedge Bw \vdash wnm & \quad \text{def of } C \\ (5) Cn_\sigma m_\sigma, w_{o\sigma\sigma} 0g_\sigma \wedge Bw \vdash w[S_n][h_{\sigma\sigma\sigma} nm] & \quad \text{def of } B, \text{ hyp., .4} \\ (6) \vdash Cn_\sigma m_\sigma \supset C[S_n][h_{\sigma\sigma\sigma} nm] & \quad \text{def of } C, .5 \end{aligned}$$

We shall need .8 as a lemma.

$$\begin{aligned} (7) \vdash Cn_\sigma m_\sigma \supset \bullet [\lambda x_\sigma \lambda y_\sigma \bullet \mathbb{N}x] 0g_\sigma \wedge B[\lambda x_\sigma \lambda y_\sigma \bullet \mathbb{N}x] \supset & \quad \text{def of } C \\ [\lambda x_\sigma \lambda y_\sigma \bullet \mathbb{N}x] nm & \\ (8) \vdash Cn_\sigma m_\sigma \supset \mathbb{N}_{o\sigma} n & .7, \text{ def of } B, 6100, 6101 \end{aligned}$$

Now we start the inductive proof with the case $n = 0$. Let $W_{o\sigma\sigma}$ be $[\lambda x_\sigma \lambda y_\sigma \bullet x = 0 \supset g_\sigma = y]$.

- (.9) $C0k_\sigma \vdash W0g_\sigma \wedge Bw \supset W0k$ def of C
- (.10) $\vdash W0g_\sigma$ 5200, def of W
- (.11) $\vdash W[Sx_\sigma][h_{\sigma\sigma\sigma}xy_\sigma] = \Box Sx = 0 \supset \Box g_\sigma = hxy$ def of W
- (.12) $\vdash W[Sx_\sigma][h_{\sigma\sigma\sigma}xy_\sigma]$ 6105, .11
- (.13) $\vdash BW$.12, def of B
- (.14) $C0k_\sigma \vdash W0k$.9, .10, .13
- (.15) $\vdash \forall k_\sigma \Box C0k \supset g_\sigma = k$.14, def of W , 5200
- (.16) $\vdash \exists m_\sigma \Box C0m \wedge \forall k_\sigma \Box C0k \supset m = k$.3, .15

Next we prove the induction part of .2.

- (.17) $.17 \vdash Cn_\sigma m_\sigma \wedge \forall k_\sigma \Box Cnk \supset \Box m = k$ (inductive) hyp
- (.18) $.17 \vdash C[Sn_\sigma][h_{\sigma\sigma\sigma}nm_\sigma]$.17, .6

We shall show that from .17 and $C[Sn_\sigma]k_\sigma$ we can prove $h_{\sigma\sigma\sigma}n_\sigma m_\sigma = k_\sigma$. So let $W_{\sigma\sigma\sigma}^2$ be $[\lambda x_\sigma \lambda y_\sigma \Box Cxy \wedge \Box x = Sn_\sigma \supset \Box hnm_\sigma = y]$.

- (.19) $C[Sn_\sigma]k_\sigma \vdash W^20g_\sigma \wedge BW^2 \supset W^2[Sn]k$ def of C
- (.20) $C[Sn_\sigma]k_\sigma \vdash W^20g_\sigma \wedge BW^2 \supset \Box hnm_\sigma = k$.19, 5200, def of W^2

Thus we wish to prove W^20g and BW^2 .

- (.21) $\vdash W^20g_\sigma$.3, 6105, def of W^2
- (.22) $W^2x_\sigma y_\sigma \vdash Cxy$ hyp, def of W^2
- (.23) $W^2x_\sigma y_\sigma \vdash C[Sx]h_{\sigma\sigma\sigma}xy$.6, .22
- (.24) $W^2x_\sigma y_\sigma \vdash Nx$.8, .22
- (.25) $.17, W^2x_\sigma y_\sigma, Sx = Sn_\sigma \vdash x = n$ hyp, 6109, .24, .17, .8
- (.26) $.17, W^2x_\sigma y_\sigma, Sx = Sn_\sigma \vdash Cny$.25, .22
- (.27) $.17, W^2x_\sigma y_\sigma, Sx = Sn_\sigma \vdash m_\sigma = y$.17, .26
- (.28) $.17, W^2x_\sigma y_\sigma, Sx = Sn_\sigma \vdash h_{\sigma\sigma\sigma}nm_\sigma = hxy$.25, .27
- (.29) $.17, W^2x_\sigma y_\sigma \vdash W^2[Sx][h_{\sigma\sigma\sigma}xy]$.23, .28, def of W^2
- (.30) $.17 \vdash BW^2$.29, def of B
- (.31) $.17 \vdash \forall k_\sigma \Box C[Sn_\sigma]k \supset \Box h_{\sigma\sigma\sigma}nm_\sigma = k$.20, .21, .30
- (.32) $.17 \vdash \exists m_\sigma \Box C[Sn_\sigma]m \wedge \forall k_\sigma \Box C[Sn]k \supset \Box m = k$.18, .31
- (.33) $\exists m_\sigma [17] \vdash [.32]$ \exists Rule: .32
- (.34) $\vdash \forall n_\sigma \exists m_\sigma \Box Cnm \wedge \forall k_\sigma \Box Cnk \supset m = k$.2, .16, .33

This completes the real work in this proof.

- (.35) $\vdash \exists_1 m_\sigma Cn_\sigma m = \exists m_\sigma \Box Cnm \wedge \forall k_\sigma \Box Cnk \supset m = k$ 5306
- (.36) $\mathbb{N}n_\sigma \vdash \exists_1 m_\sigma Cnm$.34, .35
- (.37) $\vdash \exists_1 m_\sigma Cn_\sigma m \supset Cn[\iota_{\sigma(\sigma\sigma)} \Box C_{\sigma\sigma\sigma}n_\sigma]$ 5311

- (.38) $\mathbb{N}n_\sigma \vdash Cn[\iota_{\sigma(\sigma\sigma)} \bullet Cn] .36, .37$
- (.39) $\mathbb{N}n_\sigma \vdash Cn[R_{\sigma\sigma\sigma(\sigma\sigma\sigma)} h_{\sigma\sigma\sigma} g_\sigma n] .1, .38$
- (.40) $\mathbb{N}n_\sigma \vdash Cnm_\sigma \supset \bullet Rh_{\sigma\sigma\sigma} g_\sigma n = m 5307, .36, .39$

Now we can derive the results we wish.

- (.41) $\vdash Rh_{\sigma\sigma\sigma} g_\sigma 0 = g .40, .3, 6100$
- (.42) $\vdash Cn_\sigma[Rh_{\sigma\sigma\sigma} g_\sigma n] \supset \bullet C[Sn][hn \bullet Rhgn] .6$
- (.43) $\mathbb{N}n_\sigma \vdash C[Sn][h_{\sigma\sigma\sigma} n \bullet Rhg_\sigma n] .39, .42$
- (.44) $\mathbb{N}n_\sigma \vdash \mathbb{N}[Sn] 6101$
- (.45) $\mathbb{N}n_\sigma \vdash Rh_{\sigma\sigma\sigma} g_\sigma [Sn] = \bullet hn \bullet Rhgn .40, .44, .43$

Our theorem follows from .41 and .45. ■

To illustrate the usefulness of R , define $+_{\sigma\sigma\sigma}$ as $R_{\sigma\sigma\sigma(\sigma\sigma\sigma)}[\lambda x_\sigma S_{\sigma\sigma}]$, and $\mathbf{A}_\sigma + \mathbf{B}_\sigma$ as $+_{\sigma\sigma\sigma}\mathbf{A}_\sigma\mathbf{B}_\sigma$. Then as an immediate consequence of 6400 we obtain $\vdash g_\sigma + 0 = g_\sigma \wedge \forall n_\sigma \bullet g_\sigma + [Sn_\sigma] = S[g_\sigma + n_\sigma]$. From this one can derive the usual properties of addition. (See the exercises below.) Similarly, $\times_{\sigma\sigma\sigma}$ can be defined as $[\lambda m_\sigma \bullet R_{\sigma\sigma\sigma(\sigma\sigma\sigma)}[\lambda x_\sigma \bullet +_{\sigma\sigma\sigma} m_\sigma]0_\sigma]$, and $\mathbf{A}_\sigma \times \mathbf{B}_\sigma$ as $\times_{\sigma\sigma\sigma}\mathbf{A}_\sigma\mathbf{B}_\sigma$. Then from 6400 one obtains $\vdash m_\sigma \times 0 = 0 \wedge \forall n_\sigma \bullet m_\sigma \times [Sn_\sigma] = m_\sigma + [m_\sigma \times n_\sigma]$.

Recall that a numerical function of 0 arguments is simply a number.

DEFINITION. The set of *primitive recursive* numerical functions (of zero or more arguments) is the intersection of all sets \mathcal{C} of numerical functions which have the following five properties:

- (a) $0 \in \mathcal{C}$.
- (b) $S \in \mathcal{C}$ (where S is the successor function).
- (c) For each k and i such that $1 \leq i \leq k$, $((\lambda n_1, \dots, n_k)n_i) \in \mathcal{C}$.
- (d) If h is an m -ary function (where $m \geq 1$) in \mathcal{C} and g_1, \dots, g_m are k -ary functions (where $k \geq 0$) in \mathcal{C} , then

$$((\lambda n_1, \dots, n_k)h(g_1(n_1, \dots, n_k), \dots, g_m(n_1, \dots, n_k))) \in \mathcal{C}.$$

(When $k = 0$, this means that $h(g_1, \dots, g_m) \in \mathcal{C}$.)

- (e) If g is a k -ary function in \mathcal{C} (where $k \geq 0$) and h is a $(k+2)$ -ary function in \mathcal{C} , and f is obtained from g and h by primitive recursion, then $f \in \mathcal{C}$.

The functions $((\lambda n_1, \dots, n_k)n_i)$ of (c) are called *projection* functions; these, together with the successor function and the number zero are called *initial functions*. The function of (d) is obtained by *composition* from g_1, \dots, g_m , and h . Thus a function is primitive recursive iff it can be obtained from the initial functions by a finite number of applications of composition and primitive recursion. Most of the numerical functions which are actually used in mathematics (outside the fields of logic and recursive function theory) are primitive recursive. Further discussion of primitive recursive functions will occur in §65.

DEFINITION. The set of *general recursive* numerical functions is the intersection of all sets \mathcal{C} of numerical functions which have properties (a) – (e) (in the definition of the primitive recursive functions), and also the following property:

- (f) If g is a regular $(k+1)$ -ary numerical function (where $k \geq 0$) in \mathcal{C} , and h is obtained from g by minimization, then $h \in \mathcal{C}$.

Clearly, every primitive recursive function is general recursive. There are general recursive functions which are not primitive recursive. Indeed, it can be shown (see [Mendelson, 1964, p. 250]) that the following equations define such a function:

$$\begin{aligned} f(n, 0) &= Sn \\ f(0, Sm) &= f(1, m) \\ f(Sn, Sm) &= f(f(n, Sm), m). \end{aligned}$$

The *partial recursive* functions are defined in the same way as the general recursive functions, except that minimization may be applied to functions which are not regular. Hence partial recursive functions may not be defined for certain arguments. General recursive functions are often referred to simply as recursive functions.

A numerical relation (or a set of natural numbers) is said to be *primitive recursive*, *general recursive*, or *partial recursive* iff its characteristic function is primitive recursive, general recursive, or partial recursive, respectively.

6401 Theorem. Every general recursive function is representable in \mathcal{Q}_0^∞ .

Proof:

- (a) The numeral 0_σ represents the number zero.

- (b) The wff $S_{\sigma\sigma}$ represents the successor function, since $\vdash S\bar{n} = \overline{n+1}$ by 5200 and the definition of $\overline{n+1}$.
- (c) The wff $[\lambda x_{\sigma}^1 \dots \lambda x_{\sigma}^k x_{\sigma}^i]$ represents the projection function $((\lambda n_1, \dots, n_k)n_i)$, since $\vdash [\lambda x_{\sigma}^1 \dots \lambda x_{\sigma}^k x_{\sigma}^i]\bar{n}_1 \dots \bar{n}_k = \bar{n}_i$ by 5207.
- (d) Suppose h is an m -ary function represented by $H_{\sigma\sigma\dots\sigma}$, and g_1, \dots, g_m are k -ary functions represented by $G_{\sigma\sigma\dots\sigma}^1, \dots, G_{\sigma\sigma\dots\sigma}^m$, respectively. We show that $[\lambda x_{\sigma}^1 \dots \lambda x_{\sigma}^k \bullet H[G^1 x^1 \dots x^k] \dots [G^m x^1 \dots x^k]]$ represents $((\lambda n_1, \dots, n_k)h(g_1(n_1, \dots, n_k), \dots, g_m(n_1, \dots, n_k)))$, which we shall call f .

$$\begin{aligned}
 (1.) & \vdash [\lambda x_{\sigma}^1 \dots x_{\sigma}^k \bullet H[G^1 x^1 \dots x^k] \dots [G^m x^1 \dots x^k]]\bar{n}_1 \dots \bar{n}_k \\
 &= H[G^1 \bar{n}_1 \dots \bar{n}_k] \dots [G^m \bar{n}_1 \dots \bar{n}_k] \quad \text{by Equality Rules and 5207} \\
 (2.i) & \vdash G^i \bar{n}_1 \dots \bar{n}_k = g_i(n_1, \dots, n_k) \quad \text{for } 1 \leq i \leq m, \text{ since } G^i \text{ represents } g_i \\
 (3.) & \vdash Hg_1(n_1, \dots, n_k) \dots g_m(n_1, \dots, n_k) \\
 &= h(g_1(n_1, \dots, n_k), \dots, g_m(n_1, \dots, n_k)) \quad \text{since } H \text{ represents } h \\
 (4.) & \vdash [\lambda x_{\sigma}^1 \dots \lambda x_{\sigma}^k \bullet H[G^1 x^1 \dots x^k] \dots [G^m x^1 \dots x^k]]\bar{n}_1 \dots \bar{n}_k \\
 &= f(n_1, \dots, n_k) \quad \text{Equality Rules: .1, .2}^1 - .2^m, .3
 \end{aligned}$$

The proof for the case where $k = 0$ is an obvious modification of the proof above.

- (e) Suppose g is a k -ary function represented by $G_{\sigma\sigma\dots\sigma}$, and h is a $(k+2)$ -ary function represented by $H_{\sigma\sigma\sigma\dots\sigma}$, and f is obtained by primitive recursion from g and h , so $f(m_1, \dots, m_k, 0) = g(m_1, \dots, m_k)$, and $f(m_1, \dots, m_k, Sn) = h(m_1, \dots, m_k, n, f(m_1, \dots, m_k, n))$. Let $F_{\sigma\sigma\sigma\dots\sigma}$ be $[\lambda x_{\sigma}^1 \dots \lambda x_{\sigma}^k \bullet R_{\sigma\sigma\sigma(\sigma\sigma)}[H_{\sigma\sigma\sigma\dots\sigma}x_{\sigma}^1 \dots x_{\sigma}^k] \bullet G_{\sigma\sigma\dots\sigma}x_{\sigma}^1 \dots x_{\sigma}^k]$. To show that F represents f , we show by induction on n that

$$\vdash F\bar{m}_1 \dots \bar{m}_k\bar{n} = f(m_1, \dots, m_k, n)$$

$$\begin{aligned}
 (5.) & \vdash F\bar{m}_1 \dots \bar{m}_k\bar{0} = R[H\bar{m}_1 \dots \bar{m}_k][G\bar{m}_1 \dots \bar{m}_k]0 \quad \text{def of } F, \bar{0} \\
 (6.) & \vdash R[H\bar{m}_1 \dots \bar{m}_k][G\bar{m}_1 \dots \bar{m}_k]0 = G\bar{m}_1 \dots \bar{m}_k \quad 6400 \\
 (7.) & \vdash G\bar{m}_1 \dots \bar{m}_k = g(m_1, \dots, m_k) \quad \text{since } G \text{ represents } g \\
 (8.) & \vdash F\bar{m}_1 \dots \bar{m}_k\bar{0} = f(m_1, \dots, m_k, 0) \quad .5, .6, .7, \text{ def of } f
 \end{aligned}$$

This completes the case where $n = 0$. Next we treat the induction step.

$$\begin{aligned}
 (9.) & \vdash F\bar{m}_1 \dots \bar{m}_k\bar{n} = \overline{f(m_1, \dots, m_k, n)} \quad \text{inductive hypothesis} \\
 (10.) & \vdash F\bar{m}_1 \dots \bar{m}_k(Sn) = R[H\bar{m}_1 \dots \bar{m}_k][G\bar{m}_1 \dots \bar{m}_k][S\bar{n}] \quad \text{def of } F, \bar{S}n \\
 (11.) & \vdash R[H\bar{m}_1 \dots \bar{m}_k][G\bar{m}_1 \dots \bar{m}_k][S\bar{n}] \\
 &= [H\bar{m}_1 \dots \bar{m}_k]\bar{n} \bullet R[H\bar{m}_1 \dots \bar{m}_k][G\bar{m}_1 \dots \bar{m}_k]\bar{n} \quad 6400, 6110 \\
 (12.) & \vdash R[H\bar{m}_1 \dots \bar{m}_k][G\bar{m}_1 \dots \bar{m}_k]\bar{n} = f(m_1, \dots, m_k, n) \quad .9, \text{ def of } F
 \end{aligned}$$

$$(13) \vdash H\overline{m_1} \dots \overline{m_k} \overline{nf(m_1, \dots, m_k, n)} = \overline{f(m_1, \dots, m_k, Sn)}$$

H represents h, def of f

$$(14) \vdash F\overline{m_1} \dots \overline{m_k} (\overline{Sn}) = \overline{f(m_1, \dots, m_k, Sn)}$$

.10, .11, .12, .13

- (f) By 6303, a function obtained from a representable regular function by minimization is representable.

Thus the set of representable functions has properties (a) - (f), and so includes the set of general recursive functions. ■

We shall see later (Theorem 7009) that every representable numerical function is recursive.

Informally, we may regard a numerical k -ary function f as *effectively computable* iff there is a finite algorithm, or set of rules, by means of which one may calculate $f(m_1, \dots, m_k)$ for any arguments m_1, \dots, m_k , provided no limits are imposed on the amount of time one may take, or the amount of storage (memory) one may use for recording intermediate results. Of course, this notion is only as precise as the notion of an algorithm. Nevertheless, it is clear that there can be only denumerably many algorithms, since each algorithm can be described by some finite sequence of letters and symbols of a fixed language (such as English, perhaps augmented with mathematical symbols) which has a finite alphabet. Hence, there are only denumerably many effectively computable functions. Since there are uncountably many numerical functions, some (one might even say most) numerical functions are not effectively computable.

It is clear that each recursive function is effectively computable. Indeed, implicit in the proof above that every recursive function is representable is a method for effectively calculating the value of any recursive function for arbitrary arguments. The assertion that every effectively computable numerical function (which is defined everywhere on the domain of natural numbers) is recursive is known as *Church's Thesis*. This thesis was proposed by Alonzo Church in [Church, 1936b]³. A number of quite different explications of the intuitive notion of algorithm have been proposed, and they have all turned out to define the same set of numerical functions, namely the recursive functions. On account of this and other evidence for its correctness, Church's Thesis has become generally accepted. (Additional discussion of Church's Thesis follows the proof of Theorem 7009 below.) Much of the significance of the results which we shall obtain in Chapter 7 is a consequence of Church's Thesis.

³See [Sieg, 1997] for an informative study of the development of Church's thesis.

Now that we have shown that in \mathcal{Q}_0^∞ the natural numbers can be defined and many numerical functions can be represented, we could go on to define the integers, the rational numbers, the real numbers, and the complex numbers, and formalize much of mathematics within \mathcal{Q}_0^∞ . We shall not actually carry out the details of this, but we suggest that the reader examine a book such as [Gleason, 1966], and note that what is done there could be formalized in \mathcal{Q}_0^∞ .

EXERCISES

Prove the following theorems:

$$\mathbf{X6400. } \dot{\forall} n_\sigma \dot{\forall} m_\sigma \mathbb{N}[h_{\sigma\sigma\sigma} n_\sigma m_\sigma] \supset \dot{\forall} g_\sigma \dot{\forall} n_\sigma \mathbb{N} \cdot R_{\sigma\sigma\sigma(\sigma\sigma\sigma)} h_{\sigma\sigma\sigma} g_\sigma n_\sigma.$$

$$\mathbf{X6401. } g_\sigma + 0 = g_\sigma$$

$$\mathbf{X6402. } \dot{\forall} n_\sigma \cdot g_\sigma + [Sn_\sigma] = S[g_\sigma + n_\sigma]$$

$$\mathbf{X6403. } \dot{\forall} g_\sigma \dot{\forall} n_\sigma \mathbb{N}[g_\sigma + n_\sigma]$$

$$\mathbf{X6404. } \dot{\forall} n_\sigma \cdot 0 + n_\sigma = n_\sigma$$

$$\mathbf{X6405. } \dot{\forall} n_\sigma \cdot m_\sigma + Sn_\sigma = Sm_\sigma + n_\sigma$$

$$\mathbf{X6406. } \dot{\forall} n_\sigma \cdot [Sm_\sigma] + n_\sigma = S[m_\sigma + n_\sigma]$$

$$\mathbf{X6407. } \dot{\forall} m_\sigma \dot{\forall} n_\sigma \cdot m_\sigma + n_\sigma = n_\sigma + m_\sigma$$

$$\mathbf{X6408. } \dot{\forall} n_\sigma \dot{\forall} p_\sigma \cdot [m_\sigma + n_\sigma] + p_\sigma = m_\sigma + [n_\sigma + p_\sigma]$$

$$\mathbf{X6409. } \dot{\forall} n_\sigma \cdot m_\sigma + n_\sigma = 0 \supset \cdot m_\sigma = 0 \wedge n_\sigma = 0$$

$$\mathbf{X6410. } \dot{\forall} m_\sigma \dot{\forall} n_\sigma \dot{\forall} p_\sigma \cdot p_\sigma + m_\sigma = p_\sigma + n_\sigma \supset m_\sigma = n_\sigma$$

$$\mathbf{X6411. } \dot{\forall} x_\sigma \dot{\forall} y_\sigma \cdot x_\sigma \leq y_\sigma = \exists z_\sigma \cdot x_\sigma + z_\sigma = y_\sigma$$

$$\mathbf{X6412. } \dot{\forall} x_\sigma \dot{\forall} z_\sigma \cdot x_\sigma \leq x_\sigma + z_\sigma$$

Define $\dot{+}_{(o(o\alpha))(o(o\alpha))(o(o\alpha))}$ as

$$[\lambda m_{o(o\alpha)} \lambda n_{o(o\alpha)} \lambda p_{o\alpha} \cdot \exists q_{o\alpha} \exists r_{o\alpha} \cdot m_{o(o\alpha)} q_{o\alpha} \wedge n_{o(o\alpha)} r_{o\alpha}$$

$$\wedge \forall x_\alpha [\sim q_{o\alpha} x_\alpha \vee \sim r_{o\alpha} x_\alpha] \wedge p_{o\alpha} = q_{o\alpha} \cup r_{o\alpha}].$$

Define $M_{o(o\alpha)} + N_{o(o\alpha)}$ to be $\dot{+}_{(o(o\alpha))(o(o\alpha))(o(o\alpha))} M_{o(o\alpha)} N_{o(o\alpha)}$. $\dot{+}$ expresses addition of arbitrary cardinal numbers.

X6413. Prove $\forall x_\sigma \forall y_\sigma \exists z_\sigma z_\sigma + y_\sigma = x_\sigma + y_\sigma$

X6414. Let $\exp(x, y) = x^y$, where $x^0 = 1$ and $x^{y+1} = x \cdot x^y$ for all natural numbers x and y . (Thus, $0^0 = 1$, but $0^y = 0$ when $y > 0$.) Find a wff which represents \exp , and prove that it does so.

X6415. The wff $R_{\sigma\sigma\sigma(\sigma\sigma\sigma)}[\lambda x_\sigma \lambda y_\sigma. S_{\sigma\sigma}. S_{\sigma\sigma} y_\sigma] \bar{5}$, which we shall call $F_{\sigma\sigma}$, represents a certain numerical function of one argument. Identify this function, and justify your answer with a proof.

§65. Primitive Recursive Functions and Relations

As preparation for Chapter 7, we need to show that certain numerical functions and relations are primitive recursive, and hence representable. For the sake of brevity, we shall let \mathcal{PRF}^n denote the set of n -ary primitive recursive functions, \mathcal{PRR}^n denote the set of n -ary primitive recursive relations, $\mathcal{PRF} = \bigcup_{n=0}^{\infty} \mathcal{PRF}^n$ (the set of primitive recursive functions), $\mathcal{PRR} = \bigcup_{n=0}^{\infty} \mathcal{PRR}^n$, and $\mathcal{PR} = \mathcal{PRF} \cup \mathcal{PRR}$. Also, we shall use \vec{n}_3 as an abbreviation for n_1, n_2, n_3 , or for (n_1, n_2, n_3) (where appropriate), \vec{x}_k as an abbreviation for x_1, \dots, x_k , or for (x_1, \dots, x_k) , etc. When $k = 0$, \vec{x}_k denotes the null sequence. From the definitions in §64 it is immediate that $0 \in \mathcal{PRF}^0$, $S \in \mathcal{PRF}^1$, $((\lambda \vec{n}_k)n_i) \in \mathcal{PRF}^k$ for $1 \leq i \leq k$, and \mathcal{PR} is closed under composition and primitive recursion. Also, $r \in \mathcal{PRR}^n$ iff $\chi_r \in \mathcal{PRF}^n$.

6500 Constant Functions. For all $k \geq 0$ and $n \geq 0$, $((\lambda \vec{x}_k)n) \in \mathcal{PRF}^k$. (When $k = 0$, $((\lambda \vec{x}_k)n)$ means n .)

Proof by induction on k : $0 \in \mathcal{PRF}^0$, and if $n \in \mathcal{PRF}^0$ then $S_n \in \mathcal{PRF}^0$ by composition, so $n \in \mathcal{PRF}^0$ for all n . (Thus all numerical functions of 0 arguments, i.e., all natural numbers, are primitive recursive, so no further discussion of \mathcal{PRF}^0 will be necessary.)

If $((\lambda \vec{x}_k)n) \in \mathcal{PRF}^k$, where $k \geq 0$, then $((\lambda \vec{x}_{k+1})n) \in \mathcal{PRF}^{k+1}$, since $((\lambda \vec{x}_{k+1})n)$ is obtained from $((\lambda \vec{x}_k)n)$ and $((\lambda \vec{x}_{k+2})x_{k+2})$ by primitive recursion. ■

We next wish to generalize the principle of composition to permit the following sort of inferences: if $g_1 \in \mathcal{PRF}^2$, $g_2 \in \mathcal{PRF}^3$, and $h \in \mathcal{PRF}^4$, then $((\lambda x_1, x_2, x_3)h(x_2, g_1(x_1, x_1), 5, g_2(x_2, 4, x_1))) \in \mathcal{PRF}^3$.

Let \mathcal{S} be a set of numerical functions of zero or more arguments. We shall define what it means to say that a k -ary numerical function f is explicitly definable from \mathcal{S} . So let $\mathcal{F}^=$ be a formulation of first order logic with equality, and let $\langle \mathbb{N}, \mathcal{J} \rangle$ be an interpretation of $\mathcal{F}^=$ with the set \mathbb{N} of natural numbers as domain of individuals, such that for each function g in \mathcal{S} there is a constant \bar{g} of appropriate type of $\mathcal{F}^=$ so that $\mathcal{J}\bar{g} = g$, and there is a k -ary function constant \bar{f} such that $\mathcal{J}\bar{f} = f$. We say that f is *explicitly definable* from \mathcal{S} iff there is a term t of $\mathcal{F}^=$ in which the only variables are the individual variables x_1, \dots, x_k , and the only constants are the constants \bar{g} denoting functions g in \mathcal{S} , such that

$$\langle \mathbb{N}, \mathcal{J} \rangle \models \forall x_1 \dots \forall x_k [\bar{f}x_1 \dots x_k = t].$$

Note: The wff above will be called an *explicit definition* of f from \mathcal{S} with respect to $\langle \mathbb{N}, \mathcal{J} \rangle$.

6501 Explicit Definitions. If $\mathcal{S} \subseteq \mathcal{PRF}$ and f is a numerical function explicitly definable from \mathcal{S} , then $f \in \mathcal{PRF}$.

Proof: We hold k and \mathcal{S} fixed, and prove the proposition for all k -ary functions f explicitly definable from \mathcal{S} . The proof is by induction on the construction of the term t in any explicit definition of f .

If t is an individual variable x_i , then the explicit definition is $\forall x_1 \dots \forall x_k [\bar{f}x_1 \dots x_k = x_i]$, so $f = ((\lambda n_1, \dots, n_k)n_i)$, which is in \mathcal{PRF} .

If t is an individual constant c , then $\mathcal{J}c$ must be some natural number m , so $f = ((\lambda n_1, \dots, n_k)m)$, which is in \mathcal{PRF} by 6500.

Suppose t has the form $\bar{g}t_1 \dots t_m$, where $m \geq 1$ and \bar{g} is an m -ary function constant, and each t_i is a term. For $1 \leq i \leq m$, let f_i be the k -ary numerical function with explicit definition $\forall x_1 \dots \forall x_k [\bar{f}_i x_1 \dots x_k = t_i]$, where the constants $\bar{f}_1, \dots, \bar{f}_m$ are distinct from one another, \bar{f} , and the constants in t . By inductive hypothesis, $f_i \in \mathcal{PRF}^k$. Clearly

$$\langle \mathbb{N}, \mathcal{J} \rangle \models \forall x_1 \dots \forall x_k [\bar{f}x_1 \dots x_k = \bar{g}(\bar{f}_1 x_1 \dots x_k) \dots (\bar{f}_m x_1 \dots x_k)]$$

(the parentheses being inserted only for the sake of readability), so $f = ((\lambda n_1, \dots, n_k)g(f_1(n_1, \dots, n_k), \dots, f_m(n_1, \dots, n_k)))$, so $f \in \mathcal{PRF}^k$ by composition. ■

6502 Generalized Primitive Recursion. Let $\mathcal{S} \subseteq \mathcal{PRF}$, and let g and h be k -ary and $(k+2)$ -ary functions, respectively, which are explicitly definable from \mathcal{S} . If for all natural numbers x_1, \dots, x_k , and y ,

$$f(x_1, \dots, x_i, 0, x_{i+1}, \dots, x_k) = g(\vec{x}_k), \text{ and}$$

$$f(x_1, \dots, x_i, S y, x_{i+1}, \dots, x_k) = h(\vec{x}_k, f(x_1, \dots, x_i, y, x_{i+1}, \dots, x_k)),$$

then $f \in \mathcal{PR}$. (Note: As a special case, i may be 0, i.e., the sequence x_1, \dots, x_i may be void. Similarly, i may be k , so that the sequence x_{i+1}, \dots, x_k is void.)

Proof: Let t be obtained from g and h by primitive recursion. g and h are in \mathcal{PR} by 6501, so $t \in \mathcal{PR}$. However,

$$f = ((\lambda x_1, \dots, x_i, y, x_{i+1}, \dots, x_k) t(\vec{x}_k, y)),$$

so f is explicitly definable from t , and so $f \in \mathcal{PR}$. ■

In applying 6502, we shall often replace g and h by their explicit definitions (so to speak). Having generalized composition and primitive recursion, we are in a position to expeditiously show that certain functions which we need are primitive recursive. We shall be somewhat informal, and combine propositions, definitions, and proofs whenever this is convenient. We shall often use 6500-6502 implicitly, without further mention of them. In general, our proofs will contain only the main ideas; other details may be supplied by the reader.

6503 + (addition) $\in \mathcal{PRF}^2$.

Proof:

$$x + 0 = x$$

$$x + (S y) = S(x + y)$$

■

6504. If $f \in \mathcal{PRF}^{k+1}$ (where $k \geq 0$), then $((\lambda \vec{x}_k, y) \sum_{z \leq y} f(\vec{x}_k, z)) \in \mathcal{PRF}^{k+1}$.

Proof:

$$\sum_{z \leq 0} f(\vec{x}_k, z) = f(\vec{x}_k, 0)$$

$$\sum_{z \leq \text{Sy}} f(\vec{x_k}, z) = f(\vec{x_k}, \text{Sy}) + \sum_{z \leq y} f(\vec{x_k}, z)$$
■

6505. \times (multiplication) $\in \mathcal{PRF}^2$.

Proof:

$$x \times 0 = 0$$

$$x \times (\text{Sy}) = x + (x \times y)$$
■

6506. If $f \in \mathcal{PRF}^{k+1}$ (where $k \geq 0$), then $((\lambda \vec{x_k}, y) \Pi_{z \leq y} f(\vec{x_k}, z)) \in \mathcal{PRF}^{k+1}$.

Proof:

$$\Pi_{z \leq 0} f(\vec{x_k}, z) = f(\vec{x_k}, 0)$$

$$\Pi_{z \leq \text{Sy}} f(\vec{x_k}, z) = f(\vec{x_k}, \text{Sy}) \times \Pi_{z \leq y} f(\vec{x_k}, z)$$
■

6507. $((\lambda x, y) x^y)$ (exponentiation) $\in \mathcal{PRF}^2$.

Proof:

$$x^0 = 1$$

$$x^{(\text{Sy})} = x \times (x^y)$$
■

Note: $0^0 = 1$ but $0^{(\text{Sy})} = 0$.

NOTATION. We may sometimes write x^y as $x \exp y$.

6508. ! (factorial) $\in \mathcal{PRF}^1$.

Proof:

$$0! = 1$$

$$(\text{Sy})! = y! \times \text{Sy}$$
■

6509. δ (predecessor) $\in \mathcal{PRF}^1$.

Proof:

$$\delta(0) = 0$$

$$\delta(\text{Sy}) = y$$
■

6510. $\dot{-}$ (modified subtraction or monus) $\in \mathcal{PRF}^2$.

Proof:

$$x \dot{-} 0 = x$$

$$x \dot{-} (\text{Sy}) = \delta(x \dot{-} y)$$
■

Note: $x \dot{-} y = 0$ iff $x \leq y$. Also, $x \dot{-} y = x - y$ if $x \geq y$.

6511. σ (signum) $\in \mathcal{PRF}^1$.

Proof:

$$\sigma(0) = 0$$

$$\sigma(\text{Sy}) = 1.$$
■

6512. $\bar{\sigma}$ (antisignum) $\in \mathcal{PRF}^1$.

Proof:

$$\bar{\sigma}(0) = 1$$

$$\bar{\sigma}(\text{Sy}) = 0$$
■

6513. $< \in \mathcal{PRR}^2$.

Proof: $\chi_<(x, y) = \sigma(y \dot{-} x)$

■

6514. $\leq \in \mathcal{PRR}^2$.

Proof: $x \leq y \equiv x < \text{Sy}$, so $\chi_{\leq}(x, y) = \chi_<(x, \text{Sy})$.

■

6515. $= \in \mathcal{P}\mathcal{R}\mathcal{R}^2$

Proof. $\chi_=(x, y) = ((\text{S}x) \dot{-} y) \times ((\text{S}y) \dot{-} x)$ ■

DEFINITION. Let \mathcal{S} be a set of numerical relations and functions, and let r be a k -ary numerical relation. Let \mathcal{F} be a formulation of first order logic, and let $\langle \mathbb{N}, \mathcal{J} \rangle$ be an interpretation of \mathcal{F} with the natural numbers as individuals, such that for each function or relation g in \mathcal{S} there is a constant \bar{g} of appropriate type of \mathcal{F} so that $\mathcal{J}\bar{g} = g$, and there is a k -ary predicate constant \bar{R} such that $\mathcal{J}\bar{R} = r$. We say that r is *explicitly definable from \mathcal{S} using propositional connectives* iff there is a quantifier-free wff \mathbf{B} of \mathcal{F} in which the only variables are the individual variables x_1, \dots, x_k , and the only constants are the constants \bar{g} denoting relations and functions in \mathcal{S} , such that $\langle \mathbb{N}, \mathcal{J} \rangle \models \forall x_1 \dots \forall x_k \cdot \bar{R}x_1 \dots x_k \equiv \mathbf{B}$.

6516. If $\mathcal{S} \subseteq \mathcal{P}\mathcal{R}$, and the numerical relation r is explicitly definable from \mathcal{S} using propositional connectives, then $r \in \mathcal{P}\mathcal{R}$.

EXAMPLE: Let $r(x, y) \equiv [x \neq 2y \wedge x + y < 100]$. $r \in \mathcal{P}\mathcal{R}$ by 6500, 6503, 6505, 6513, 6515, and 6516.

Proof: We hold k and \mathcal{S} fixed, and prove the proposition for all k -ary relations r explicitly definable from \mathcal{S} using propositional connectives. The proof is by induction on the construction of a wff \mathbf{B} explicitly defining r from \mathcal{S} . We may assume that \mathbf{B} contains no propositional connectives other than \sim and \vee .

If \mathbf{B} is an atomic wff $\bar{P}t_1 \dots t_n$, and p is the relation in \mathcal{S} such that $\mathcal{J}\bar{P} = p$, $\langle \mathbb{N}, \mathcal{J} \rangle \models \forall x_1 \dots \forall x_k [\bar{P}x_1 \dots x_k \equiv \bar{P}t_1 \dots t_n]$, so $\langle \mathbb{N}, \mathcal{J} \rangle \models \forall x_1 \dots \forall x_k [\bar{\chi}_r x_1 \dots x_k = \bar{\chi}_p t_1 \dots t_n]$, so χ_r is explicitly definable from χ_p and functions in \mathcal{S} , so $\chi_r \in \mathcal{P}\mathcal{R}\mathcal{F}$ by 6501. Hence $r \in \mathcal{P}\mathcal{R}\mathcal{R}$.

Suppose \mathbf{B} has the form $\sim \mathbf{C}$. Let p be the k -ary numerical relation with explicit definition $\forall x_1 \dots \forall x_k [\bar{P}x_1 \dots x_k \equiv \mathbf{C}]$, where the predicate constant \bar{P} does not occur in \mathbf{C} . $p \in \mathcal{P}\mathcal{R}\mathcal{R}$ by inductive hypothesis, and for all natural numbers \vec{n}_k , $r(\vec{n}_k)$ iff not $p(\vec{n}_k)$. Thus $\chi_r(\vec{n}_k) = \bar{\sigma}(\chi_p(\vec{n}_k))$, so $r \in \mathcal{P}\mathcal{R}\mathcal{R}$.

Suppose \mathbf{B} has the form $[\mathbf{C} \vee \mathbf{D}]$. Let p and q be the k -ary numerical relations with explicit definitions $\forall x_1 \dots \forall x_k [\bar{P}x_1 \dots x_k \equiv \mathbf{C}]$ and $\forall x_1 \dots \forall x_k [\bar{Q}x_1 \dots x_k \equiv \mathbf{D}]$, respectively, where \bar{P} and \bar{Q} do not occur in $[\mathbf{C} \vee \mathbf{D}]$. By inductive hypothesis p and q are in $\mathcal{P}\mathcal{R}\mathcal{R}$, and for all natural numbers \vec{n}_k , $r(\vec{n}_k)$ iff $p(\vec{n}_k)$ or $q(\vec{n}_k)$. Thus $\chi_r(\vec{n}_k) = \sigma(\chi_p(\vec{n}_k) + \chi_q(\vec{n}_k))$, so $r \in \mathcal{P}\mathcal{R}\mathcal{R}$. ■

We shall use the expression $(\exists z < y)B$ in our meta-language to mean $\exists z[z < y \wedge B]$, i.e., there is a number z less than y such that B . Similarly, $(\forall z < y)B$ means $\forall z[z < y \supset B]$. Analogous conventions apply when $<$ is replaced by \leq . Note that $(\exists z < 0)B$ is false, and $(\forall z < 0)B$ is true.

6517 Bounded Quantification. If $f \in \mathcal{PRF}^k$ (where $k \geq 0$) and $r \in \mathcal{PRR}^{k+1}$, then the following are in \mathcal{PRR}^k :

- (a) $(\lambda \vec{x}_k)(\exists z \leq f(\vec{x}_k))r(\vec{x}_k, z)$
- (b) $(\lambda \vec{x}_k)(\exists z < f(\vec{x}_k))r(\vec{x}_k, z)$
- (c) $(\lambda \vec{x}_k)(\forall z \leq f(\vec{x}_k))r(\vec{x}_k, z)$
- (d) $(\lambda \vec{x}_k)(\forall z < f(\vec{x}_k))r(\vec{x}_k, z)$

Proof:

- (a) $\chi_{(a)} = (\lambda \vec{x}_k)\sigma(\sum_{z \leq f(\vec{x}_k)} \chi_r(\vec{x}_k, z)).$
- (b) $(\exists z < f(\vec{x}_k, z))r(\vec{x}_k, z) \equiv (\exists z \leq f(\vec{x}_k))[z \neq f(\vec{x}_k) \wedge r(\vec{x}_k, z)].$

By (a), 6501, 6515, and 6516, this is in \mathcal{PRR} .

- (c) $\chi_{(c)} = (\lambda \vec{x}_k)\Pi_{z \leq f(\vec{x}_k)} \chi_r(\vec{x}_k, z).$
- (d) $(\forall z < f(\vec{x}_k))r(\vec{x}_k, z) \equiv (\forall z \leq f(\vec{x}_k))[z = f(\vec{x}_k) \vee r(\vec{x}_k, z)].$ ■

We shall use the expression $(\mu z \leq y)B$ in our meta-language to denote the least number $z \leq y$ such that B if such a z exists; if no such z exists, we let $(\mu z \leq y)B = Sy$. Also, we shall use the expression $(\text{Max } z \leq y)B$ in our meta-language to denote the largest number $z \leq y$ such that B if such a z exists; if no such z exists, we let $(\text{Max } z \leq y)B = 0$.

6518 Bounded Minimization and Maximization. If $f \in \mathcal{PRF}^k$ and $r \in \mathcal{PRR}^{k+1}$ (where $k \geq 0$), then the following are in \mathcal{PRF}^k :

- (a) $(\lambda \vec{x}_k)(\mu z \leq f(\vec{x}_k))r(\vec{x}_k, z)$
- (b) $(\lambda \vec{x}_k)(\text{Max } z \leq f(\vec{x}_k))r(\vec{x}_k, z)$

Proof:

- (a) Let $p = (\lambda \vec{x}_k, z)(\forall y \leq z) \sim r(\vec{x}_k, y)$. $p \in \mathcal{PRR}$ by 6516 and 6517, and $(\mu z \leq f(\vec{x}_k))r(\vec{x}_k, z) = \sum_{z \leq f(\vec{x}_k)} \chi_p(\vec{x}_k, z)$. (See the example in Figure 6.1.)

	y	0	1	2	3	4	5	6	7	...
	$r(y)$	F	F	F	F	T	F	T	T	...
$py \equiv (\forall w \leq y) \sim r(w)$		T	T	T	T	F	F	F	F	...
	$\chi_p(y)$	1	1	1	1	0	0	0	0	...
	$\sum_{z \leq y} \chi_p(z)$	1	2	3	4	4	4	4	4	...

Figure 6.1: $(\mu z \leq y)r(z)$

(b) Let $g(\vec{x}, w) = w \dashv (\mu y \leq w)r(\vec{x}, w \dashv y)$. Clearly $g \in \mathcal{PRF}^{k+1}$. We shall show that (1) $g(\vec{x}, w) = (\text{Max } z \leq w)r(\vec{x}, z)$. From this the desired result follows.

Note that if there is no $z \leq w$ such that $r(\vec{x}, z)$, then $g(\vec{x}, w) = w \dashv \text{Sw} = 0$, as desired. If such a z exists, let z_0 be the largest one. Thus (2) $r(\vec{x}, z_0)$, (3) $z_0 \leq w$, and (4) $\forall z[z_0 < z \leq w \Rightarrow \sim r(\vec{x}, z)]$.

Let $y_0 = w - z_0$. We shall show that (5) $y_0 = (\mu y \leq w)r(\vec{x}, w - y)$. By (2), $r(\vec{x}, w \dashv y_0)$. Suppose $0 \leq y < y_0$. Then $w \geq w - y > w - y_0 = z_0$ so $\sim r(\vec{x}, w \dashv y)$ by (4). This establishes (5).

Now $g(\vec{x}, w) = w \dashv y_0$ (by (5) and the definition of g) = z_0 (by the definition of y_0) = $(\text{Max } z \leq w)r(\vec{x}, z)$ (by the definition of z_0). ■

REMARK. It is important to note the distinction between bounded minimization, which can be used to define primitive recursive functions, and unbounded minimization, which can be used to define general recursive functions.

6519 Definition by Cases. Let $g_1, \dots, g_n \in \mathcal{PRF}^k$ and $r_1, \dots, r_n \in \mathcal{PRR}^k$, and suppose the relations r_1, \dots, r_n are mutually exclusive and exhaustive, i.e., for all $\vec{x}_k \in \mathbb{N}^k$, there is exactly one r_i such that $r_i(\vec{x}_k)$. Let $f(\vec{x}_k) = g_i(\vec{x}_k)$ iff $r_i(\vec{x}_k)$ for $1 \leq i \leq n$. Then $f \in \mathcal{PRF}^k$.

Proof: $f = (\lambda \vec{x}_k)((g_1(\vec{x}_k) \times \chi_{r_1}(\vec{x}_k)) + \dots + (g_n(\vec{x}_k) \times \chi_{r_n}(\vec{x}_k)))$. ■

6520. $((\lambda x, y)x \text{ Divides } y) \in \mathcal{PRR}^2$

Proof: $x \text{ Divides } y \equiv (\exists z \leq y)[x \times z = y]$. ■

6521. Let $\text{Prime}(x)$ iff x is a prime number. Then $\text{Prime} \in \mathcal{PRR}^1$.

Proof:

$$\text{Prime}(x) \equiv \neg 2 \leq x \wedge (\forall y \leq x)[y \text{ Divides } x \supset \neg y = x \vee y = 1] \quad \blacksquare$$

6522. Let $\text{Pr}(n)$ be the n th odd prime number, where $\text{Pr}(0) = 2$. Then $\text{Pr} \in \mathcal{PRF}^1$.

Proof:

$$\text{Pr}(0) = 2$$

$$\text{Pr}(\text{Sn}) = (\lambda m \leq \text{S}(\text{Pr}(n)!))[\text{Prime}(m) \wedge \text{Pr}(n) < m]$$

To verify the correctness of the recursion equation, we must show that if (by inductive hypothesis) $\text{Pr}(n)$ is prime, then there is a prime number m such that $\text{Pr}(n) < m \leq \text{S}(\text{Pr}(n)!)$. (The proof is essentially Euclid's proof that there are infinitely many primes.) Let $b = \text{Pr}(n)!$. If there is a prime p such that $\text{Pr}(n) < p \leq b$, we are done, so suppose there is no such prime. In this case we show that Sb is prime. Let p be any prime less than Sb . Then $p \leq \text{Pr}(n)$, so p divides $\text{Pr}(n)!$, which is b . But $p \geq 2$, so p cannot also divide Sb . Thus no prime less than Sb divides Sb , so Sb is prime.

Let $h = (\lambda x, y)(\mu m \leq \text{S}(y!))[\text{Prime}(m) \wedge y < m]$. $h \in \mathcal{PRF}^2$ by 6521, 6513, 6516, 6508, and 6518, and $\text{Pr}(\text{Sn}) = h(n, \text{Pr}(n))$, so $\text{Pr} \in \mathcal{PRF}^1$ by primitive recursion. \blacksquare

Here are a few values for Pr :

n	0	1	2	3	4	5	6	7	8	9	10	11	12
$\text{Pr}(n)$	2	3	5	7	11	13	17	19	23	29	31	37	41

Note: $n < \text{Pr}(n)$. This is readily established by induction.

We wish to use positive integers to represent finite sequences of natural numbers, and to do this we shall make extensive use of the fundamental fact from number theory that every positive integer can be factored uniquely into a product of primes. We shall associate with each positive integer the sequence of natural numbers which constitute the exponents in its prime factorization. Of course, all but finitely many of the terms of this sequence are zero. Thus, with the number 84, which is $2^2 \times 3^1 \times 5^0 \times 7^1$, we associate the sequence $(2, 1, 0, 1, 0, 0, 0, \dots)$, or the finite sequence $(2, 1, 0, 1)$.

6523. Let $n \mathcal{G}\ell x$ be the exponent of $\text{Pr}(n)$ in the prime factorization of x , with $n \mathcal{G}\ell 0 = 0$. Then $\mathcal{G}\ell \in \mathcal{PRF}^2$.

Note: If $x = \prod_{i=0}^{\infty} \text{Pr}(i)^{a_i}$, then $n \mathcal{G}\ell x = a_n$, the n th term in the sequence $a_0, a_1, a_2, a_3, \dots$ associated with x . The name “ $\mathcal{G}\ell$ ” for this function was introduced in [Gödel 1931]. “Glied” means “term” in German.

Proof: $n \mathcal{G}\ell x = (\text{Max } y \leq x) [\text{Pr}(n)^y \text{ Divides } x]$. There is always some number y such that $\text{Pr}(n)^y$ divides x , since $\text{Pr}(n)^0 = 1$. Since $\text{Pr}(n) \geq 2$, it is readily established by induction that $y < \text{Pr}(n)^y$, so if $y = n \mathcal{G}\ell x$ and $x \neq 0$, then $y < \text{Pr}(n)^y \leq x$, so the given bound is adequate. ■

Note: $n \mathcal{G}\ell 0 = 0 = n \mathcal{G}\ell 1$ for all n . Also, $n \mathcal{G}\ell x < x$ if $x \neq 0$.

EXAMPLE: $0 \mathcal{G}\ell 84 = 2$, $1 \mathcal{G}\ell 84 = 1$, $2 \mathcal{G}\ell 84 = 0$, $3 \mathcal{G}\ell 84 = 1$, $n \mathcal{G}\ell 84 = 0$ for $n > 3$.

6524. Let $E(x)$ (the extent of the sequence represented by x) be the largest number n such that $\text{Pr}(n)$ divides x , with $E(0) = 0 = E(1)$. Then $E \in \mathcal{PRF}^1$.

Proof: $E(x) = (\text{Max } n \leq x) [\text{Pr}(n) \text{ Divides } x]$. ■

Clearly if $\text{Pr}(n)$ divides x , then $n < \text{Pr}(n) \leq x$, so $E(x) \leq x$.

Note: $x = \prod_{i \leq E(x)} \text{Pr}(i)^i \mathcal{G}\ell x$ if $x \neq 0$.

Note: $E(\prod_{i \leq n} \text{Pr}(i)^{a_i}) = n$ if $a_n \neq 0$.

We next define a concatenation function $*$ such that if x and y represent the sequences (a_0, \dots, a_n) and (b_0, \dots, b_m) , respectively, then $x * y$ represents the sequence $(a_0, \dots, a_n, b_0, \dots, b_m)$. We do this in such a way that the numbers 0 and 1 can be used to represent the empty sequence.

6525. Let $x * y = x \times \prod_{i \leq E(y)} \text{Pr}(E(x) + 1 + i)^i \mathcal{G}\ell y$ if $1 < x$
 $= y$ if $x \leq 1$.

Then $* \in \mathcal{PRF}^2$.

Proof: By 6519. ■

Note that if $x \leq 1$, then $x * y = y$, but if $x > 1$ and $y \leq 1$, then $x * y = x$ (since $E(y) = 0$ and $0 \mathcal{G}\ell y = 0$). Also, if $a_n \neq 0$ and $b_m \neq 0$,

$$(\text{Pr}(0)^{a_0} \times \dots \times \text{Pr}(n)^{a_n}) * (\text{Pr}(0)^{b_0} \times \dots \times \text{Pr}(m)^{b_m}) = \\ \text{Pr}(0)^{a_0} \times \dots \times \text{Pr}(n)^{a_n} \times \text{Pr}(n+1)^{b_0} \times \dots \times \text{Pr}(n+1+m)^{b_m}.$$

Hence, $*$ is associative, i.e., for all natural numbers x, y , and z , $x * (y * z) = (x * y) * z$.

EXAMPLE: $21 * 300 = (2^0 \times 3^1 \times 5^0 \times 7^1) * (2^2 \times 3^1 \times 5^2) = 2^0 \times 3^1 \times 5^0 \times 7^1 \times 11^2 \times 13^1 \times 17^2 = 9,546,537$.

Since the binary numerical functions $+$, $*$, and \times are all associative, we shall often omit the parentheses associated with them. In contexts where two or three of these functions occur together, we shall use the convention that $+$ has the largest scope, and \times the smallest. Thus, in restoring parentheses in an expression involving $+$, $*$, and \times , first restore parentheses associated with \times , giving them the smallest possible scope, and then those associated with $*$, giving them the smallest possible scope, and then (if necessary) those associated with $+$. For example, $a + b \times c * d * e + f \times g \times h + i * j$ stands for $a + (b \times c) * d * e + (f \times g \times h) + i * j$, and for $a + ((b \times c) * d * e) + (f \times g \times h) + (i * j)$.

6526. Let $\text{Lg} = (\lambda x)(\mu y \leq x)(\exists z \leq x)[x = z^y \wedge \text{Prime}(z)]$. Then $\text{Lg} \in \mathcal{PRF}^1$.

EXAMPLE: $\text{Lg}(16) = 4$ since $16 = 2^4$.

Chapter 7

Incompleteness and Undecidability

Logical systems such as \mathcal{Q}_0^∞ have infinitely many theorems—indeed, infinitely many essentially different theorems—but these theorems are finitely generated. There may be infinitely many axioms, but a finite description in the meta-language suffices to specify them. Wffs are finite, and proofs are finite. Such features seem to be inescapable characteristics of logical systems which can actually be used by finite, mortal men.

It is natural to ask whether such systems are inherently incapable of dealing adequately with the infinite variety and complexity of the mathematical world. One's response will obviously depend on one's notions of adequacy, but a hasty general affirmative answer may be deterred by considering the possibility of a fine balance between the powers of asking and answering questions in or about a suitable logical system, so that unanswerable questions cannot even be asked. It is natural to ask whether there is an ultimately satisfactory logical system, in which one can express the important and generally useful ideas which arise naturally in the study of the logical system itself as well as in other parts of mathematics, and in which one can prove those truths which can be expressed, but no sentences which are not true.

The results in this chapter, which involve some of the most important and profound ideas in the history of logic, should illuminate these questions. The results are formulated in terms of \mathcal{Q}_0 and its extensions, but the ideas behind them are of very wide applicability. With appropriate modifications, the results about \mathcal{Q}_0^∞ apply to any logical system for formalizing mathematics which is not too weak, inadequately specified, inconsistent, or in some other

way unsatisfactory. The significance of these results is particularly apparent in the context of Hilbert's program. In the early part of the twentieth century, the great mathematician David Hilbert set forth a set of goals for research in logic known as Hilbert's program . As applied to some particular mathematical theory, the goals of the program included:

- (1) Define a formal system \mathcal{S} such that the theorems of \mathcal{S} are precisely the true statements of the theory.
- (2) Prove the consistency of \mathcal{S} by elementary methods whose correctness could not reasonably be doubted.
- (3) Solve the decision problem for \mathcal{S} . Thus all mathematical questions in the theory would in principle be answered, and further research would simply be devoted to improving the efficiency of the decision procedure.

As we shall see, it was soon discovered that these goals are unattainable for many mathematical theories (such as the theory of natural numbers, the theory of real numbers, and extensions of them).

Without waiting for the precise definitions of the concepts involved, we now give a brief preview of the main results in this chapter. \mathcal{Q}_0^∞ is essentially incomplete, i.e., it is incomplete and has no consistent complete extension which is precisely specifiable (recursively axiomatizable). Consequently, there is no recursively axiomatizable extension of \mathcal{Q}_0 whose theorems are precisely the wffs valid in all standard models. The consistency of \mathcal{Q}_0^∞ is not provable by methods which can be formalized in \mathcal{Q}_0^∞ . The decision problems for \mathcal{Q}_0 and \mathcal{Q}_0^∞ are unsolvable. On the purely semantic side, truth for \mathcal{Q}_0^∞ is not definable in \mathcal{Q}_0^∞ , i.e., the set of sentences of \mathcal{Q}_0^∞ which are true (in a given model) is not definable by any wff of \mathcal{Q}_0^∞ .

The crucial role of Church's Thesis in this chapter should be noted. An effective operation corresponds, via a suitable encoding, to a numerical function which, by Church's Thesis, must be recursive, and hence is representable in \mathcal{Q}_0^∞ . Thus, if we accept Church's Thesis and find an operation (such as deciding which wffs are theorems of \mathcal{Q}_0) which corresponds to a function which is not representable in \mathcal{Q}_0^∞ , then we must accept the fact that no algorithm can execute that operation.

§70. Gödel Numbering

We wish to be able to express certain statements about \mathcal{Q}_0^∞ by wffs of \mathcal{Q}_0^∞ . For this purpose, we assign numbers to the type symbols, primitive symbols, formulas, and sequences of formulas of \mathcal{Q}_0 , and use the numerals of \mathcal{Q}_0^∞ denoting these numbers as names for the corresponding syntactic entities. The results we shall obtain will not really depend on our particular choice of a numbering system, since the basic ideas we shall use can be applied to obtain these results for a rather wide variety of numbering systems. Most of the numbering systems which naturally come to mind can be translated into one another by primitive recursive functions. We shall make extensive use of the results and terminology of §65.

DEFINITION. To each type symbol α we assign a natural number “ α ” as follows by induction on the construction of α :

- (1) “ o ” = 3
- (2) “ i ” = 5
- (3) “ $(\beta\gamma)$ ” = $3^{“\beta”} \times 5^{“\gamma”}$

Note that if α and β are distinct type symbols, then “ α ” ≠ “ β ”.

7000 Theorem. Let $\text{TypeSymbol} = ((\lambda n)n \text{ is the number of some type symbol})$. Then $\text{TypeSymbol} \in \mathcal{PRR}^1$.

Proof: In order to determine whether a nonzero number S_n is the number of a type symbol, one may need to know whether certain numbers $k < S_n$ are numbers of type symbols; however, information about n may not suffice. Hence we cannot define $\chi_{\text{TypeSymbol}}$ directly by primitive recursion. Instead we define a function f such that from $f(n)$ we can determine for each number $k \leq n$ whether $\text{TypeSymbol}(k)$.

Let $f = ((\lambda n)\Pi_{i \leq n}\text{Pr}(i)^{\chi_{\text{TypeSymbol}}(i)})$. We first show that $f \in \mathcal{PRF}^1$ by 6502.

$$f(0) = 1 \quad (\text{since } 1 = 2^0 = 2^{\chi_{\text{TypeSymbol}}(0)})$$

$$f(S_n) = f(n) \times [(\text{Pr}(S_n))\exp(\chi_=(S_n, 3) + \chi_=(S_n, 5))$$

$$+ [\chi_=(S_n, 3^{\text{G}\ell(S_n)} \times 5^{\text{G}\ell(S_n)})]$$

$$\times (1 \text{ G}\ell S_n \text{ G}\ell f(n) \times (2 \text{ G}\ell S_n \text{ G}\ell f(n)))]$$

since $\text{TypeSymbol}(\text{Sn})$ iff $\text{Sn} = "o" = 3$ or $\text{Sn} = "i" = 5$ or there exist a and b such that $\text{Sn} = 3^a \times 5^b$ (in which case $a = 1 \not\in \text{Sn}$ and $b = 2 \not\in \text{Sn}$) and $\text{TypeSymbol}(a)$ and $\text{TypeSymbol}(b)$, i.e., $a \not\in f(n) = \chi_{\text{TypeSymbol}}(a) = 1$ and $b \not\in f(n) = 1$. Now $\chi_{\text{TypeSymbol}} = ((\lambda n)n \not\in f(n)) \in \mathcal{PRF}$. ■

DEFINITION. To each primitive symbol \mathbf{Z} of \mathcal{Q}_0 we assign a number ‘ \mathbf{Z} ’ as follows:

\mathbf{Z}	[]	λ	$\iota_{(\iota(o\alpha))}$	$Q_{((o\alpha)\alpha)}$	f_α	g_α	h_α	...
‘ \mathbf{Z} ’	7	11	13	17 “ $(\iota(o\alpha))$ ”	19 “ $((o\alpha)\alpha)$ ”	23 “ α ”	29 “ α ”	31 “ α ”	...
x_α	y_α	nth variable x_α of type α							
103 “ α ”	107 “ α ”	$\Pr(7 + n) " \alpha "$							

To each formula (sequence of primitive symbols) $\mathbf{Z}_0 \dots \mathbf{Z}_n$, we assign the number “ $\mathbf{Z}_0 \dots \mathbf{Z}_n$ ” = $\Pi_{i \leq n} \Pr(i) " \mathbf{Z}_i "$.

Note that if $\mathbf{Z}_0 \dots \mathbf{Z}_n$ is a sequence of primitive symbols of which $\mathbf{W}_0 \dots \mathbf{W}_m$ is a subsequence, then “ $\mathbf{W}_0 \dots \mathbf{W}_m$ ” \leq “ $\mathbf{Z}_0 \dots \mathbf{Z}_n$ ”.

To each sequence $\mathbf{A}_0, \dots, \mathbf{A}_n$ of formulas, we assign the number “ $\mathbf{A}_0, \dots, \mathbf{A}_n$ ” = $\Pi_{i \leq n} \Pr(i) " \mathbf{A}_i "$.

7001 Theorem. Let $Q = (\lambda n)19 \exp((3 \exp(3^3 \times 5^n)) \times 5^n)$. Then $Q \in \mathcal{PRF}^1$, and for each type symbol α , $Q(" \alpha ") = 'Q_{o\alpha\alpha}'$.

7002 Theorem. Let $\text{Variable} = ((\lambda n)n$ is the number of a variable of \mathcal{Q}_0). Then $\text{Variable} \in \mathcal{PRR}^1$.

Proof: $\text{Variable}(n) \equiv (\exists a \leq n)(\exists i \leq n)[8 \leq i \wedge \text{TypeSymbol}(a) \wedge n = \Pr(i)^a]$. ■

7003 Theorem. Let $\text{ProperSymbol} = ((\lambda n)n$ is the number of a variable or constant of \mathcal{Q}_0). Then $\text{ProperSymbol} \in \mathcal{PRR}^1$.

Proof $\text{ProperSymbol}(n) \equiv [\text{Variable}(n) \vee n = \iota_{(\iota(o\alpha))}] \vee (\exists a \leq n)[\text{TypeSymbol}(a) \wedge n = Q(a)]$. ■

7004 Theorem. Let $\text{Wff} = ((\lambda x)$ there is a wff \mathbf{A}_α of \mathcal{Q}_0 such that

$x = "A_\alpha"$).

$$\begin{aligned}\text{Let } \text{Type}(x) &= "\alpha" \text{ if } x = "A_\alpha"; \\ &= 0 \text{ if not } \text{Wff}(x).\end{aligned}$$

Then $\text{Wff} \in \mathcal{PRR}^1$ and $\text{Type} \in \mathcal{PRF}^1$.

Proof: $\chi_{\text{Wff}} = ((\lambda x)\sigma(\text{Type}(x)))$, so it suffices to show that $\text{Type} \in \mathcal{PRF}^1$. We can determine $\text{Type}(x)$ from x and $\text{Type}(k)$ for certain numbers $k < x$. Therefore we let $f(x) = \Pi_{i \leq x} \text{Pr}(i)^{\text{Type}(i)}$, and show that f can be obtained by primitive recursion from primitive recursive functions. This will establish the theorem, since $\text{Type}(x) = x \mathcal{G}\ell f(x)$.

First we make some definitions. The reader may wish to postpone studying the details of these definitions until he has surveyed the rest of the proof. The reader should also review the definition of the function $*$ (see 6525), and observe that if \mathbf{A} and \mathbf{B} are formulas, then " $\mathbf{A} * \mathbf{B}$ " = " \mathbf{AB} ".

Let $A = (\lambda y)(\exists x \leq y)[\text{ProperSymbol}(x) \wedge y = 2^x]$; $\mathbf{A}(y)$ means that there is a proper symbol Z_a such that $y = 2^{Z_a}$ = " Z_a ". If this is the case, " α " = $\text{Lg}'(Z_a)$ = $\text{Lg}(\text{Lg}(y))$.

Let $B = (\lambda w, i, j, y)[j \mathcal{G}\ell w \neq 0 \wedge i \mathcal{G}\ell w = 3^1 \mathcal{G}\ell(i \mathcal{G}\ell w) \times 5^j \mathcal{G}\ell w \wedge y = 2^7 * i * j * 2^{11}]$. Suppose $w = f(n)$ and $i \leq n$ and $j \leq n$ and $\mathbf{B}(w, i, j, y)$. Then $\text{Type}(j) = j \mathcal{G}\ell w \neq 0$, so there is a wff \mathbf{B}_β such that $j = "\mathbf{B}_\beta"$, and $j \mathcal{G}\ell w = "\beta"$. Also, there is a number a such that $\text{Type}(i) = i \mathcal{G}\ell w = 3^a \times 5^{\beta}$, so there is a wff $\mathbf{A}_{\alpha\beta}$ such that $i = "\mathbf{A}_{\alpha\beta}"$ and $y = [" * "\mathbf{A}_{\alpha\beta} * "\mathbf{B}_\beta * "] = "[\mathbf{A}_{\alpha\beta}\mathbf{B}_\beta]"$. This wff has type α , and " α " = $1 \mathcal{G}\ell ("(\alpha\beta)") = 1 \mathcal{G}\ell(i \mathcal{G}\ell w)$.

Let $C = (\lambda w, i, j, y)[\text{Variable}(j) \wedge i \mathcal{G}\ell w \neq 0 \wedge y = (2^7 \times 3^{13} \times 5^j) * i * 2^{11}]$. Suppose $w = f(n)$ and $i \leq n$ and $j \leq n$ and $\mathbf{C}(w, i, j, y)$. Then there is a variable \mathbf{x}_β such that $j = "\mathbf{x}_\beta"$, so " β " = $\text{Lg}(j)$. Also, $\text{Type}(i) = i \mathcal{G}\ell w \neq 0$, so there is a wff \mathbf{A}_α such that $i = "\mathbf{A}_\alpha"$, and " α " = $i \mathcal{G}\ell w$. Thus $y = "[\lambda \mathbf{x}_\beta * "\mathbf{A}_\alpha * "] = "[\lambda \mathbf{x}_\beta \mathbf{A}_\alpha]"$. This wff has type $(\alpha\beta)$, and " $(\alpha\beta)$ " = $3^{\alpha} \times 5^{\beta} = 3^i \mathcal{G}\ell w \times 5^{\text{Lg}(j)}$.

$$f(0) = 2^0 = 1. \text{ Also,}$$

$$f(Sn) = f(n) \times [\text{Pr}(Sn) \exp([\text{Lg}(\text{Lg}(Sn)) \times \chi_A(Sn)])$$

$$+ \sum_{i \leq n} \sum_{j \leq n} ([(1 \mathcal{G}\ell(i \mathcal{G}\ell f(n))) \times \chi_B(f(n), i, j, Sn)]$$

$$+ [3^i \mathcal{G}\ell f(n) \times 5^{\text{Lg}(j)} \times \chi_C(f(n), i, j, Sn)])]).$$

To see this, note that $f(Sn) = f(n) \times \text{Pr}(Sn)^{\text{Type}(Sn)}$. $\text{Type}(Sn) = 0$ unless Sn is the number of a wff, in which case exactly one of the following conditions must occur:

- (1) $Sn = "Z_\alpha"$ for some primitive symbol Z_α , so $A("Z_\alpha")$ and $\text{Type}(Sn) = "\alpha"$.
- (2) $Sn = "[A_{\alpha\beta} B_\beta]"$ for some wffs $A_{\alpha\beta}$ and B_β , so $B(f(n), "A_{\alpha\beta}", "B_\beta", "[A_{\alpha\beta} B_\beta]"),$ and $\text{Type}(Sn) = "\alpha"$.
- (3) $Sn = "[\lambda x_\beta A_\alpha]"$ for some wff A_α and variable x_β , so $C(f(n), "A_\alpha", "x_\beta", "[\lambda x_\beta A_\alpha]"),$ and $\text{Type}(Sn) = "\alpha\beta"$.

Since $A, B, C \in \mathcal{PRR}$, we see that $f \in \mathcal{PRF}$. ■

7005 Theorem. Let $\text{Axiom} = ((\lambda n) \text{ there is an axiom } A \text{ of } Q_0^\infty \text{ such that } n = "A")$. Then $\text{Axiom} \in \mathcal{PRR}^1$.

Proof: For $i = 2, 3, 4_1, 4_2, 4_3, 4_4, 4_5$, let $\text{Axiom}^i = ((\lambda x)x \text{ is the number of an instance of Axiom Schema } i)$. We show that each of these is in \mathcal{PRR} .

$\text{Axiom } 2^\alpha$ is $[[\supset_{ooo} [[Q_{o\alpha\alpha} x_\alpha] y_\alpha]] [[Q_{ooo} [h_{o\alpha} x_\alpha]] [h_{o\alpha} y_\alpha]]]$.

Also, ' $h_{o\alpha}$ ' = $31 \exp(3^3 \times 5^\alpha)$, ' x_α ' = 103^α , and ' y_α ' = 107^α . Hence

$$\begin{aligned} \text{Axiom}^2(n) \equiv & (\exists a \leq n)(\text{TypeSymbol}(a) \wedge n = "[[\supset_{ooo} [" * 2^{Q(a)} \times \\ & 3^{103 \exp a} \times 5^{11} \times 7^{107 \exp a} * "]] [[Q_{ooo} [" * 2^{31 \exp(3^3 \times 5^a)} \times \\ & 3^{103 \exp a} \times 5^{11} \times 7^{11} \times 11^7 \times 13^{31 \exp(3^3 \times 5^a)} \times \\ & 17^{107 \exp a} * "]]])], \end{aligned}$$

so $\text{Axiom}^2 \in \mathcal{PRR}$.

$\text{Axiom } 3^{\alpha\beta}$ is

$$[[Q_{ooo} [[Q_{o(\alpha\beta)(\alpha\beta)} f_{\alpha\beta}] g_{\alpha\beta}]] [[Q_{o(o\beta)(o\beta)} [\lambda x_\beta T_o]] [\lambda x_\beta [[Q_{o\alpha\alpha} [f_{\alpha\beta} x_\beta]] [g_{\alpha\beta} x_\beta]]]]]]].$$

The proof that $\text{Axiom}^3 \in \mathcal{PRR}$ is similar to that for Axiom^2 .

$\text{Axiom Schema } 4_1$ has the form $[[Q_{o\beta\beta} [[\lambda x_\alpha B_\beta] A_\alpha]] B_\beta]$, where B_β is a primitive constant or variable distinct from x_α . Hence

$$\begin{aligned} \text{Axiom}^{4_1}(n) \equiv & (\exists x \leq n)(\exists b \leq n)(\exists a \leq n)(\text{Variable}(x) \wedge \text{ProperSymbol}(b) \\ & \wedge b \neq x \wedge \text{Wff}(a) \wedge \text{Type}(a) = \text{Lg}(x) \wedge n = 2^7 \times 3^7 \\ & \times 5^{\text{Q}(\text{Lg}(b))} \times 7^7 \times 11^7 \times 13^{13} \times 17^x \times 19^b \times 23^{11} * a * \\ & 2^{11} \times 3^{11} \times 5^b \times 7^{11}), \end{aligned}$$

so Axiom^{4₁} ∈ PRR¹.

Axiom Schema 4₂ has the form [[Q_{oαα}[[λx_αx_α]A_α]]A_α], so

$$\begin{aligned} \text{Axiom}^{4_2}(n) \equiv & (\exists x \leq n)(\exists a \leq n)(\text{Variable}(x) \wedge \text{Wff}(a) \wedge \text{Type}(a) = \text{Lg}(x) \\ & \wedge n = 2^7 \times 3^7 \times 5^{\text{Q}(\text{Lg}(x))} \times 7^7 \times 11^7 \times 13^{13} \times 17^x \times 19^x \\ & \times 23^{11} * a * 2^{11} \times 3^{11} * a * 2^{11}), \end{aligned}$$

so Axiom^{4₂} ∈ PRR.

Axiom 4₃ has the form

$$[[Q_{o\beta\beta}[[\lambda x_\alpha[B_{\beta\gamma}C_\gamma]]A_\alpha]][[[\lambda x_\alpha B_{\beta\gamma}]A_\alpha][[\lambda x_\alpha C_\gamma]A_\alpha]]].$$

We leave further details to the reader.

Axiom Schema 4₄ has the form

$$[[Q_{o(\delta\gamma)(\delta\gamma)}[[\lambda x_\alpha[\lambda y_\gamma B_\delta]]A_\alpha]][[\lambda y_\gamma[[\lambda x_\alpha B_\delta]A_\alpha]]]$$

where y_γ is distinct from x_α and all variables in A_α.

$$\begin{aligned} \text{Axiom}^{4_4}(n) \equiv & (\exists x \leq n)(\exists y \leq n)(\exists a \leq n)(\exists b \leq n)(\text{Variable}(x) \wedge \\ & \text{Variable}(y) \wedge \text{Wff}(b) \wedge \text{Wff}(a) \wedge \text{Type}(a) = \text{Lg}(x) \wedge y \\ & \neq x \wedge (\forall i \leq E(a))[y \neq i \text{ } \mathcal{G}\ell \text{ } a] \wedge n = “[“ * \\ & (2 \exp Q(3^{\text{Type}(b)} \times 5^{\text{Lg}(y)})) * “[“ * [\lambda” * 2^x * “[“ * \lambda” * 2^y * b * “]]”]” \\ & * a * “[“][\lambda” * 2^y * “[“ * [\lambda” * 2^x * b * 2^{11} * a * “]]”]”). \end{aligned}$$

Thus Axiom^{4₄} ∈ PRR¹.

¹The phrase ‘Wff(a)’ above is actually redundant, but we leave it in as an aid to intelligibility.

Axiom Schema 4₅ has the form $[[Q_{o(\delta\alpha)(\delta\alpha)}[[\lambda x_\alpha[\lambda x_\alpha B_\delta]]A_\alpha]][\lambda x_\alpha B_\delta]].$ We leave further details concerning this axiom to the reader.

$$\begin{aligned} \text{Axiom}(n) \equiv & (n = \text{"Axiom 1"} \vee \text{Axiom}^2(n) \vee \text{Axiom}^3(n) \vee \text{Axiom}^{4_1}(n) \\ & \vee \text{Axiom}^{4_2}(n) \vee \text{Axiom}^{4_3}(n) \vee \text{Axiom}^{4_4}(n) \vee \text{Axiom}^{4_5}(n) \\ & \vee n = \text{"Axiom 5"} \vee n = \text{"Axiom 6"}). \end{aligned}$$

Thus Axiom $\in \mathcal{P}\mathcal{R}\mathcal{R}^1.$

■

7006 Theorem Let Rule = $(\lambda m, n, p)$ [there exist wffs A_α , B_α , C_o , and D_o such that $m = "[A_\alpha = B_\alpha]"$ and $n = "C_o"$ and $p = "D_o"$, and D_o is obtained from $[A_\alpha = B_\alpha]$ and C_o by Rule R]. Then Rule $\in \mathcal{P}\mathcal{R}\mathcal{R}^3.$

Proof:

$$\begin{aligned} \text{Rule}(m, n, p) \equiv & (\exists a \leq m)(\exists b \leq m)[\text{Wff}(a) \wedge \text{Wff}(b) \wedge \text{Type}(a) = \text{Type}(b) \\ & \wedge m = 2^7 \times 3^7 \times 5^{\text{Q}(\text{Type}(a))} * a * 2^{11} * b * 2^{11} \wedge \text{Type}(n) \\ & = 3 \wedge \text{Type}(p) = 3 \wedge (\exists j \leq n)(\exists k \leq n)[n = j * a * k \wedge \\ & p = j * b * k \wedge (\text{E}(j)) \text{ G}\ell j \neq 13]]. \end{aligned}$$

To see this, note that if Rule(m, n, p) is true then $m = "[[Q_{o\alpha\alpha}A_\alpha]B_\alpha]",$ and there exist formulas (which need not be wffs) J and K such that $C = JA_\alpha K$ and $D = JB_\alpha K.$ Of course, J and K may be empty sequences of symbols, in which case they have 1 as number. Also, the last symbol of J must not be a $\lambda.$

■

DEFINITION. A logistic system \mathcal{L}^2 is a *pure extension* of a logistic system \mathcal{L}^1 iff the systems have the same wffs and rules of inference, and every axiom of \mathcal{L}^1 is a theorem of $\mathcal{L}^2.$ (Hence every theorem of \mathcal{L}^1 is a theorem of $\mathcal{L}^2).$

DEFINITION. A logistic system \mathcal{L} is *[primitive] recursively axiomatized* iff its set of axioms is [primitive] recursive (with respect to a suitable Gödel numbering). Note that since there are primitive recursive translations between most numbering systems, this definition does not really depend on the particular numbering system that is used.

It is a consequence of Church's Thesis that a system is recursively axiomatized iff there is an effective test for determining of an arbitrary wff whether it is an axiom. Thus we may assume that any system which will actually be used (by mortal men) to prove theorems will be recursively axiomatized.

We have shown in Theorem 7005 that \mathcal{Q}_0^∞ is primitive recursively axiomatized, and hence is a (primitive) recursively axiomatized pure extension of \mathcal{Q}_0 .

7007 Theorem. Let \mathcal{A} be any [primitive] recursively axiomatized pure extension of \mathcal{Q}_0 . Let $\text{Proof}^{\mathcal{A}} = (\lambda m, n) [\text{There is a proof } \mathbf{P}^0, \dots, \mathbf{P}^k \text{ in } \mathcal{A} \text{ such that } m = ``\mathbf{P}^k'' \text{ and } n = ``\mathbf{P}^0, \dots, \mathbf{P}^k'']$. Then $\text{Proof}^{\mathcal{A}}$ is a [primitive] recursive relation.

Proof:

$$\begin{aligned} \text{Proof}^{\mathcal{A}}(m, n) \equiv & [m = E(n) \text{ gl } n \wedge (\forall k \leq E(n))[\text{Axiom}^{\mathcal{A}}(k \text{ gl } n) \\ & \vee (\exists i < k)(\exists j < k)\text{Rule}(i \text{ gl } n, j \text{ gl } n, k \text{ gl } n)]]. \end{aligned}$$

A quick review of our results about primitive recursive functions and relations will show that they also apply to recursive functions and relations, so if $\text{Axiom}^{\mathcal{A}}$ is recursive, $\text{Proof}^{\mathcal{A}}$ is also. ■

7008 Theorem. Let $\text{Num} = ((\lambda n)``\bar{n}")$. Then $\text{Num} \in \mathcal{PRF}^1$.

Proof:

$$\text{Num}(0) = ``0_\sigma"$$

$$\text{Num}(S_n) = 2^7 * ``S_{\sigma\sigma}" * \text{Num}(n) * 2^{11} \text{ (since } \overline{S_n} = [S_{\sigma\sigma}\bar{n}]).$$

7009 Theorem. If \mathcal{A} is a consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ , then every numerical function and every numerical relation representable in \mathcal{A} is recursive.

Proof: By 6113 it suffices to prove the theorem for functions. Let f be an n -ary numerical function (where we may assume $n \geq 1$), and let $\mathbf{F}_{\sigma\sigma\dots\sigma}$ be a wff which represents f . Let us sketch the main ideas of the proof below before getting into the technical details. Given numbers $\vec{x_n}$, we can find $f(\vec{x_n})$ looking for a proof of a theorem of the form $\mathbf{F}\vec{x_1}\dots\vec{x_n} = \bar{z}$. This involves searching simultaneously for z and the proof, but we can instead search for a number y of the form $2^z \times 3^p$, where p is the number of the proof. When we consider the least such number y for a fixed choice of $\vec{x_n}$,

we have $z = 0 \text{ G}\ell y = 0 \text{ G}\ell ((\mu y)\text{Proof}^{\mathcal{A}}(\text{“F}\bar{x}_1 \dots \bar{x}_n = \bar{z}”, p))$. Of course, y depends on the numbers \bar{x}_n . We shall define a numerical relation S_n so that $S_n(\text{“F”}, \bar{x}_n, 2^z \times 3^p)$ is true iff p is the number of a proof of $\text{F}\bar{x}_1 \dots \bar{x}_n = \bar{z}$.

Now let us start the formal argument by defining $S_n = (\lambda j, \bar{x}_n, y)[\text{Proof}^{\mathcal{A}}(2^7 \times 3^7 \times 5^{‘Q_{o\sigma\sigma}’} \times (\text{Pr}(3))^7 \times \dots \times (\text{Pr}(2+n))^7 * j * \text{Num}(x_1) * 2^{11} * \text{Num}(x_2) * 2^{11} * \dots * \text{Num}(x_n) * 2^{11} * 2^{11} * \text{Num}(0 \text{ G}\ell y) * 2^{11}, 1 \text{ G}\ell y)]$. We assert that $(\lambda \bar{x}_n, y)S_n(\text{“F”}, \bar{x}_n, y)$ is regular, and that $f = (\lambda \bar{x}_n)0 \text{ G}\ell ((\mu y)S_n(\text{“F”}, \bar{x}_n, y))$.

Given arbitrary natural numbers \bar{x}_n , let $z = f(\bar{x}_n)$. Since $\text{F}_{\sigma\sigma\dots\sigma}$ represents f , there is a proof in \mathcal{A} of the wff $\text{F}_{\sigma\sigma\dots\sigma}\bar{x}_1 \dots \bar{x}_n = \bar{z}$, i.e., $[[Q_{o\sigma\sigma}[\dots[[\text{F}\bar{x}_1]\bar{x}_2]\dots\bar{x}_n]]\bar{z}]$. Let p be the smallest possible number of such a proof, and $y = 2^z \times 3^p$. Hence $\text{Proof}^{\mathcal{A}}(\text{“F}\bar{x}_1 \dots \bar{x}_n = \bar{z}”, p)$ is true. Note that

$$\begin{aligned} S_n(\text{“F”}, \bar{x}_n, y) &= \text{Proof}^{\mathcal{A}}(\text{“}[Q_{o\sigma\sigma}[\dots[“] * “\text{F}” * “\bar{x}_1” * “]”} \\ &\quad * “\bar{x}_2” * “]” * \dots * “\bar{x}_n” * “]” * “]” * “\bar{z}” * “]”, p) \\ &= \text{Proof}^{\mathcal{A}}(\text{“}[Q_{o\sigma\sigma}[\dots[[\text{F}\bar{x}_1]\bar{x}_2]\dots\bar{x}_n]]\bar{z}”}, p). \end{aligned}$$

Thus $S_n(\text{“F”}, \bar{x}_n, y)$ is true, and the relation is regular. Moreover, since \mathcal{A} is consistent, by 6112 there is no number $w \neq z$ such that $\vdash_{\mathcal{A}} \text{F}\bar{x}_1 \dots \bar{x}_n = \bar{w}$, so y is the least number such that $S_n(\text{“F”}, \bar{x}_n, y)$. Hence $f(\bar{x}_n) = z = 0 \text{ G}\ell y = 0 \text{ G}\ell ((\mu y)S_n(\text{“F”}, \bar{x}_n, y))$.

Since $(\lambda \bar{x}_n, y)S_n(\text{“F”}, \bar{x}_n, y)$ is regular and recursive, f is recursive. ■

REMARK. Theorem 7009 also applies to extensions of \mathcal{G}_0^∞ which are not pure extensions, i.e., extensions with additional primitive constants. One simply modifies the Gödel numbering (and hence the proof) to accommodate the additional constants.

7010 Theorem. If \mathcal{A} is any consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ , then a numerical function is representable in \mathcal{A} if and only if it is recursive.

Proof: By 6401 and 7009.

Thus the representable functions are precisely the recursive functions in a wide variety of extensions of \mathcal{Q}_0^∞ , some of which are significantly stronger than \mathcal{Q}_0^∞ . Actually, the situation is much more general than this. In a wide variety of logical systems in which number theory can be formalized, the representable functions are precisely the recursive functions.

Of course, if a formalization of number theory is quite weak, one may not be able to represent all recursive functions in it. However, there are finitely axiomatized formulations of number theory in first-order logic in which all recursive functions are representable. On the other hand, it is clear that every numerical function is representable in any inconsistent system.

It seems intuitively clear that if a function is computable, then an algorithm for computing its values can be expressed in some formal language. Hence, if that language is sufficiently strong, the function will actually be representable in it. On the other hand, if a function f is representable (in a consistent language where there is an effective test for deciding whether a given wff is an axiom), it must be computable. For given numbers \vec{x}_n , one can effectively find $f(\vec{x}_n)$ simply by enumerating the theorems until one finds one of the form $\mathbf{F}\vec{x}_1 \dots \vec{x}_n = \vec{z}$ (where \mathbf{F} represents f); one then knows that $f(\vec{x}_n) = z$. (Of course, this is the intuitive idea behind the proof of 7009.) Thus it seems clear that in sufficiently rich languages, the representable functions are precisely the computable functions. The fact that the representable functions are the recursive functions in a wide variety of languages, which differ among themselves in many other respects, shows that the class of recursive functions is a very natural class, and provides further evidence for Church's Thesis that it is the class of computable functions.

Before concluding this section, we note that the proof of 7009 provides certain additional information about recursive functions.

7011 Normal Form Theorem. For each natural number $n \geq 1$ there is a relation $S_n \in \mathcal{P}\mathcal{R}\mathcal{R}^{n+2}$ such that for each n -ary recursive function f , there is a number j such that $(\lambda\vec{x}_n, y)S_n(j, \vec{x}_n, y)$ is regular, and

$$f = (\lambda\vec{x}_n)0 \text{ } \mathcal{G}\ell ((\mu y)S_n(j, \vec{x}_n, y)).$$

Proof: Let \mathcal{A} be \mathcal{Q}_0^∞ in the proof of 7009. \mathcal{Q}_0^∞ can be shown to be consistent (by semantic methods), and by 7005 and 7007 $\text{Proof}^{\mathcal{Q}_0^\infty} \in \mathcal{P}\mathcal{R}\mathcal{R}$, so $S_n \in \mathcal{P}\mathcal{R}\mathcal{R}^{n+2}$. ■

Note that Theorem 7011 provides an enumeration (with repetitions) of the n -ary recursive functions. Of course, some numbers do not correspond to any function in this enumeration.

7012 Theorem. Every recursive function can be defined with at most one application of minimization.

Proof: By 7011. ■

Parable. There was once a king who was quite jealous of the scholars who taught the royal school, since he did not like to be surrounded by people smarter than himself. He decided to give them a test in arithmetic, and to behead the scholars who finished it. He insisted that the scholars all work diligently on the test for four hours each day, omit no details, and take no shortcuts. The first scholar quickly did the problems using the usual algorithm for addition that is taught in grade school, left the palace, and was executed. The second scholar defined addition recursively as in 6503, carried out all the steps of the computation, and worked several days before completing the test and meeting his fate. This third scholar defined $+_{\sigma\sigma\sigma}$ as $R_{\sigma\sigma\sigma(\sigma\sigma\sigma)}[\lambda x_\sigma S_{\sigma\sigma}]$ and performed all additions by giving complete proofs of the corresponding theorems in \mathcal{Q}_0^∞ without using any derived rules of inference. He lived to a ripe old age. The fourth scholar defined $+$ as $\lambda x_1 \lambda x_2 \ 0 \text{ } \mathcal{G}\ell ((\mu y)S_2(" +_{\sigma\sigma\sigma}" , x_1, x_2, y))$. His descendants are still working on the test, and have inherited the kingdom.

EXERCISES

X7000. Show that Axiom⁴⁵ $\in \mathcal{PRR}$.

X7001. Construct an example of the test the king might have given the scholars in the parable above, make reasonable assumptions about computational methods used by the scholars and their rates of work, and compute the time they would actually take to complete the test.

X7002. A set of natural numbers is said to be *recursively enumerable* (r.e.) iff it is empty or is the range of a recursive function, and a set \mathcal{S} of formulas is said to be *recursively enumerable* iff $\{"\mathbf{A}" \mid \mathbf{A} \in \mathcal{S}\}$ is a recursively enumerable set of numbers. Prove that if \mathcal{A} is any recursively axiomatized extension of \mathcal{Q}_0 , then the set of theorems of \mathcal{A} is recursively enumerable.

§71. Gödel's Incompleteness Theorems

DEFINITIONS. Let \mathcal{A} be any extension of \mathcal{Q}_0 .

- (1) \mathcal{A} is *consistent* iff not $\vdash_{\mathcal{A}} F_o$.
- (2) \mathcal{A} is ω -*consistent* iff there is no closed wff \mathbf{B}_{oo} of \mathcal{A} such that $\vdash_{\mathcal{A}} \exists y_\sigma \mathbf{B}_{oo} y_\sigma$ and for each natural number k , $\vdash_{\mathcal{A}} \sim \mathbf{B}_{oo} \bar{k}$.
- (3) \mathcal{A} is *complete* iff for each sentence \mathbf{A}_o of \mathcal{A} , $\vdash_{\mathcal{A}} \mathbf{A}_o$ or $\vdash_{\mathcal{A}} \sim \mathbf{A}_o$.
- (4) \mathcal{A} is ω -*complete* iff $\vdash_{\mathcal{A}} \forall y_\sigma \mathbf{B}_{oo} y_\sigma$ for each closed wff \mathbf{B}_{oo} of \mathcal{A} such that $\vdash_{\mathcal{A}} \mathbf{B}_{oo} \bar{k}$ for each natural number k .

7100 Proposition. If \mathcal{A} is an ω -consistent extension of \mathcal{Q}_0 , then \mathcal{A} is consistent.

Proof: If \mathcal{A} is inconsistent, then every wff_o is a theorem of \mathcal{A} , so \mathcal{A} is ω -inconsistent. ■

One of the most basic and ancient paradoxes, which was well known to the ancient Greeks², is the *liar paradox*, which is invoked by the statement “This statement is false”. There are various ways in which one can construct a statement which refers to itself, as was illustrated by Example 3 near the beginning of §10. When one considers whether the statement is true or false, the paradox becomes evident. If it is false, then the assertion that the statement is false is itself false, so the statement is true. On the other hand, if it is true, then the assertion it makes, which is that the statement is false, is true, so the statement is false.

From this one may conclude that informal reasoning in natural language can lead to a contradiction, and one is led to ask what happens when one tries to imitate such an argument in a formal language. It turns out that this can lead to profound insights into the formal language.

The liar paradox, as expressed above, concerns truth, which is a semantic concept. In §73 we shall consider the consequences of the semantic form of the liar paradox for extensions of \mathcal{Q}_0^∞ , but first let us note that the paradox can be cast in a syntactic form if we replace the statement above by “This statement is not provable”. If one has a complete axiomatization of the set of sentences true in some particular model, truth in that model should correspond to provability in the theory, so the paradox should still apply, but in a syntactic form. This idea leads to the proof of Gödel's famous Incompleteness Theorem, which we discuss next.

The theorem shows that no ω -consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ can be complete. Moreover, an explicit example of a sentence which is neither provable nor refutable is provided. Clearly any satisfactory formalization of number theory must be ω -consistent. Nevertheless, in Rosser's extension (7200) of Gödel's Theorem, we shall replace the requirement of ω -consistency by the simple requirement of consistency.

By 6401, 6525, 7007, and 7008 we have established that if \mathcal{A} is any recursively axiomatized pure extension of \mathcal{Q}_0^∞ , then the functions * and Num, and the relation Proof^A, are representable in \mathcal{Q}_0^∞ , and hence in \mathcal{A} . Let $*_{\sigma\sigma\sigma}$, $Num_{\sigma\sigma}$ and $Proof_{\sigma\sigma}^A$ be wffs of \mathcal{Q}_0^∞ which represent them. Sometimes we will omit the type symbols from these wffs. The context will indicate

²See [Kleene, 1952, p. 39]

whether we are referring to the function $*$ or the wff which represents it. Note that if \mathbf{A} and \mathbf{B} are formulas, $\vdash \overline{\text{"A}} *_{\sigma\sigma\sigma} \overline{\text{"B}} = \overline{\text{"AB"}}$.

7101 Gödel's Incompleteness Theorem [Gödel, 1931]. Let \mathcal{A} be any recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Let $\star_{o\sigma}$ be the wff

$$[\lambda x_\sigma \dot{\forall} y_\sigma \sim \text{Proof}^{\mathcal{A}}[\overline{128} * x_\sigma * [\text{Num}_{\sigma\sigma} x_\sigma] * \overline{2048}]y_\sigma].$$

Let $m = \star_{o\sigma}$ and let $\star\star_o$ be $[\star_{o\sigma} \overline{m}]$.

- (1) If \mathcal{A} is consistent, then not $\vdash_{\mathcal{A}} \star\star$.
- (2) If \mathcal{A} is ω -consistent, then not $\vdash_{\mathcal{A}} \sim \star\star$.

Proof. We shall write $\vdash_{\mathcal{A}} \mathbf{B}$ simply as $\vdash \mathbf{B}$.

- (.1) $\vdash \overline{128} * \overline{m} * [\text{Num}_{\sigma\sigma} \overline{m}] * \overline{2048} = \overline{\star\star}$
since the wffs $*$ and Num represent the corresponding functions, and
 $128 * m * \text{Num}(m) * 2048 = "[" * \star_{o\sigma} * "\overline{m}" * "]" = "[\star_{o\sigma} \overline{m}] = \star\star$.
- (.2) $\vdash \star\star = \dot{\forall} y_\sigma \sim \text{Proof}^{\mathcal{A}} \overline{\star\star} y_\sigma$ λ , Def of $\star\star$; R: .1

Now suppose $\vdash \star\star$. Then

$$(.) \vdash \dot{\forall} y_\sigma \sim \text{Proof}^{\mathcal{A}} \overline{\star\star} y_\sigma \quad .2$$

Let n be the number of a proof of $\star\star$.

- (.4) $\vdash \text{Proof}^{\mathcal{A}} \overline{\star\star} \overline{n}$
since the wff Proof represents the numerical relation Proof.
- (.5) $\vdash \sim \text{Proof}^{\mathcal{A}} \overline{\star\star} \overline{n}$ VI: .3
- (.6) $\vdash F_o$ $.4, .5$

Thus if $\vdash \star\star$, then \mathcal{A} is inconsistent.

Next suppose \mathcal{A} is ω -consistent and $\vdash \sim \star\star$.

$$(.) \vdash \dot{\exists} y_\sigma \text{Proof}^{\mathcal{A}} \overline{\star\star} y_\sigma \quad .2, 6000$$

\mathcal{A} is consistent by 7100, so not $\vdash \star\star$, so for each natural number k ,

$$(.) \vdash \sim \text{Proof}^{\mathcal{A}} \overline{\star\star} \overline{k}$$

since k is not the number of a proof of $\star\star$, and the wff *Proof* represents the relation Proof. But this contradicts the ω -consistency of \mathcal{A} . Hence if \mathcal{A} is ω -consistent, not $\vdash \sim \star\star$. ■

Note from 7101.2 that with respect to our Gödel numbering, $\star\star$ does indeed express the assertion “I am not provable”. The reader should study the definitions of $\star_{o\sigma}$ and $\star\star_{o\sigma}$ carefully in order to understand how they permit 7101.2 to be proved. $\star\star$ is a sentence about natural numbers which is true in all standard models of \mathcal{A} but is not provable in \mathcal{A} . Of course, if one ignores the Gödel numbering, and seeks to understand $\star\star$ directly as a statement about natural numbers, one finds that it is rather complicated. For an example of a true but unprovable statement about natural numbers having a somewhat simpler form, see [Davis *et al.*, 1976, pp. 345-350].

7102 Corollary. Every consistent recursively axiomatized pure extension \mathcal{A} of \mathcal{Q}_0^∞ is ω -incomplete.

Proof: The wff $\star\star$ of 7101 is not provable in \mathcal{A} , so (as in 7101.8) for every natural number k , $\vdash \sim [\overline{\text{Proof}_{o\sigma\sigma}^{\mathcal{A}} \text{“}\star\star\text{”}}]k$. However, the wff $\forall y_\sigma \sim [\overline{\text{Proof}_{o\sigma\sigma}^{\mathcal{A}} \text{“}\star\star\text{”}}]y_\sigma$ is not provable in \mathcal{A} , since it is equivalent (by 7101.2) to $\star\star$. ■

We postpone further discussion of the significance of Gödel's Incompleteness Theorem until we eliminate the condition of ω -consistency in §72.

Henceforth in this section we let \mathcal{A} be any recursively axiomatized pure extension of \mathcal{Q}_0^∞ , and let $\vdash \mathbf{B}$ mean $\vdash_{\mathcal{A}} \mathbf{B}$.

We will next establish Gödel's Second Incompleteness Theorem by two arguments, which are presented in the proofs of Theorem 7110 and Corollary 7115 below. The theorem says that if \mathcal{A} is consistent, then it cannot be proved consistent by methods which can be formalized within \mathcal{A} . This implies, of course, that the consistency of \mathcal{A} can be proved within \mathcal{A} if and only if \mathcal{A} is inconsistent! The first argument essentially consists of translating from our meta-language into \mathcal{A} the proof of the first part of 7101. Thus no considerations of ω -consistency arise. As is customary, we shall not carry out all the details of the formalization. The facts we use without detailed proof are expressed in Lemmas 7104 and 7106 below.

DEFINITION. Let *Theorem* _{$o\sigma$} be the wff $[\lambda x_\sigma \exists y_\sigma \overline{\text{Proof}_{o\sigma\sigma}^{\mathcal{A}} x_\sigma y_\sigma}]$.

7103 Lemma. If $\vdash \mathbf{A}$, then $\vdash \text{Theorem}_{o\sigma} \overline{\text{“A”}}$.

Proof: If $\vdash \mathbf{A}$, then \mathbf{A} has a proof; let p be the number of such a proof. Hence the numerical relation $\text{Proof}^{\mathcal{A}}(\text{“A”}, p)$ holds, so

$$\begin{aligned} &\vdash \text{Proof}_{o\sigma\sigma}^{\mathcal{A}} \overline{\text{“A”}} \overline{p} \\ &\vdash \dot{\exists} y_{\sigma} \text{Proof}_{o\sigma\sigma}^{\mathcal{A}} \overline{\text{“A”}} y_{\sigma} && 6110, \exists \text{ Gen} \\ &\vdash \text{Theorem}_{o\sigma} \overline{\text{“A”}} && \lambda \end{aligned}$$

■

It should be remarked that in spite of Lemma 7103, the wff $\text{Theorem}_{o\sigma}$ does not represent the set of numbers of theorems of \mathcal{A} . In fact, we shall see (7300) that this set is not representable. Indeed, if we could prove $\vdash \sim \text{Theorem} \overline{\text{“A”}_o}$ for some wff \mathbf{A}_o , we could prove within \mathcal{A} the consistency of \mathcal{A} .

7104 Lemma. $\vdash \dot{\forall} x_{\sigma} \bullet \text{Theorem}_{o\sigma} x_{\sigma} \supset \text{Theorem}_{o\sigma} [\overline{[\text{Theorem}_{o\sigma}] * \text{Num } x_{\sigma} * \overline{2048}}]$

This theorem expresses 7103 within \mathcal{A} . The proof of 7103 (which depends heavily on our results about representability) is purely syntactic, and so can be formalized within \mathcal{Q}_0^∞ , and hence in \mathcal{A} . The result of this formalization is a proof of 7104. We omit the lengthy details.

In order to see that 7104 is a correct formalization of 7103, it is perhaps easiest to consider the following corollary:

7105 Corollary. For any wff_o \mathbf{A} ,

$$\vdash \text{Theorem}_{o\sigma} \overline{\text{“A”}} \supset \text{Theorem}_{o\sigma} \overline{[\text{Theorem}_{o\sigma} \overline{\text{“A”}}]}.$$

Proof:

$$(.) \vdash \text{Theorem} \overline{\text{“A”}} \supset \bullet \text{Theorem} \overline{[\text{Theorem} \overline{[\text{Theorem} * \text{Num} \overline{\text{“A”}} * \overline{2048}]}]}$$

7104

If $n = \text{“A”}$, then $[\text{Theorem}] * \text{Num}(n) * \overline{2048} = [\text{Theorem}] * \overline{n} * \overline{[}] = [\text{Theorem} \overline{\text{“A”}}]$, so

$$(.) \vdash \overline{[\text{Theorem}]} * \text{Num} \overline{\text{“A”}} * \overline{2048} = \overline{[\text{Theorem} \overline{\text{“A”}}]}$$

since $*$ and Num represent the corresponding functions.

7105 follows by Rule R. ■

7106 Lemma. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \blacksquare \text{Theorem}_{oo}[\overline{“[[Q_{ooo}] * x_\sigma * 2048 * y_\sigma * 2048]} \supset \blacksquare \text{Theorem}_{oo}x_\sigma \supset \blacksquare \text{Theorem}_{oo}y_\sigma]$.

This lemma expresses within \mathcal{A} the fact that if $\vdash A = B$ and $\vdash A$, then $\vdash B$, which we established in 5201.

We again omit the details of the proof. Actually the proof of this lemma, though long, is quite straightforward, since wffs representing the various numerical functions and relations we have defined can be obtained by a rather direct translation of our definitions into Q_0 . The proof would include the following steps:

- (.1) $.1 \vdash \mathbb{N}x_\sigma \wedge \mathbb{N}y_\sigma \wedge \text{Theorem}[\overline{“[[Q_{ooo}] * x_\sigma * 2048 * y_\sigma * 2048]} \wedge \text{Theorem } x_\sigma]$ hyp
- (.2) $.1 \vdash \text{Type } x_\sigma = \overline{3} \wedge \text{Type } y_\sigma = \overline{3}$
- (.3) $.1 \vdash \text{Rule}[\overline{“[[Q_{ooo}] * x_\sigma * 2048 * y_\sigma * 2048]x_\sigma y_\sigma”}$ from the definition of Rule
- (.4) $.1 \vdash \text{Theorem } y_\sigma$ from the definition of Proof

Actually, if one wishes to write out all the details, it may be slightly easier to prove directly the corollaries below, which are all we really need from 7106.

7107 Corollary.

$\vdash \text{Theorem}_{oo}\overline{“A_o = B_o”} \supset \blacksquare \text{Theorem}_{oo}\overline{“A_o”} \supset \text{Theorem}_{oo}\overline{“B_o”}$.

Proof:

- (.1) $\vdash \text{Theorem}_{oo}[\overline{“[[Q_{ooo}] * \overline{A} * 2048 * \overline{B} * 2048]”} \supset \blacksquare \text{Theorem}_{oo}\overline{“A”} \supset \text{Theorem}_{oo}\overline{“B”}]$ 7106
- (.2) $\vdash \overline{“[[Q_{ooo}] * \overline{A} * 2048 * \overline{B} * 2048]”} = \overline{“[[Q_{ooo}A]B]”}$ since the wff * represents the function *.

Then use Rule R to complete the proof. ■

7108 Corollary. If $\vdash A = B$, then $\vdash \text{Theorem}_{oo}\overline{“A”} \supset \text{Theorem}_{oo}\overline{“B”}$.

Proof:

- (.1) $\vdash A = B$ by hypothesis
- (.2) $\vdash \text{Theorem}\overline{“A = B”}$ 7103: .1
- (.3) $\vdash \text{Theorem}\overline{“A”} \supset \text{Theorem}\overline{“B”}$ MP: .2, 7107

DEFINITION. Consis_o^A is the wff $\sim \text{Theorem}_{oo}^A \overline{\text{``}F_o\text{''}}$.

We shall often write Consis_o^A simply as Consis .

7109 Corollary. For any wff A , $\vdash \text{Theorem}_{oo} \overline{\text{``}A\text{''}} \wedge \text{Theorem}_{oo} \overline{\text{``}\sim A\text{''}} \supset \sim \text{Consis}$.

Proof:

- (.1) $\vdash \sim A = \bullet A = F_o$ Rule P
 - (.2) $\vdash \text{Theorem} \overline{\text{``}\sim A\text{''}} \supset \text{Theorem} \overline{\text{``}A_o = F_o\text{''}}$ 7108: .1
 - (.3) $\vdash \text{Theorem} \overline{\text{``}A_o = F_o\text{''}} \supset \bullet \text{Theorem} \overline{\text{``}A\text{''}} \supset \text{Theorem} \overline{\text{``}F_o\text{''}}$ 7107
 - (.4) $\vdash \text{Theorem} \overline{\text{``}A\text{''}} \wedge \text{Theorem} \overline{\text{``}\sim A\text{''}} \supset \sim \text{Consis}$
- Rule P: .2, .3; def of Consis
-

7110 Gödel's Second Incompleteness Theorem [Gödel, 1931]. Let \mathcal{A} be any recursively axiomatized pure extension of \mathcal{Q}_0^∞ . If \mathcal{A} is consistent, then not $\vdash_{\mathcal{A}} \text{Consis}_o^A$.

Proof: We first restate the proof of the first part of 7101 using our present terminology. Let \star_{oo} be $[\lambda x_\sigma \sim \text{Theorem}_{oo}[\overline{128} * x_\sigma * [\text{Num}_{oo} x_\sigma] * \overline{2048}]]$. (Note that this is a different wff from that defined in the proof of 7101, though the two wffs are equivalent.) Let $m = \star_{oo}$ and let $\star\star_o$ be $[\star_{oo} \overline{m}]$.

- (a) $\vdash \text{Theorem} \overline{\text{``}\star\star\text{''}} = \sim \star\star$ by the definition of $\star\star$, and the fact that the wffs $*$ and Num represent the functions $*$ and Num (see 7101.1).
- (b) Suppose $\vdash \star\star$. Then
- (c) $\vdash \text{Theorem} \overline{\text{``}\star\star\text{''}}$ 7103: b
- (d) $\vdash \sim \star\star$ R: a, c
- (e) Hence by (b) and (d), \mathcal{A} is inconsistent.
- (f) Hence, if \mathcal{A} is consistent, not $\vdash \star\star$.

Next we formalize the argument above:

- (b1) $b1 \vdash \text{Theorem} \overline{\text{``}\star\star\text{''}}$ hyp
- (c1) $b1 \vdash \text{Theorem} \overline{\text{``}[\text{Theorem} \overline{\text{``}\star\star\text{''}}]\text{''}}$ MP: b1, 7105
- (d1) $\vdash \text{Theorem} \overline{\text{``}[\text{Theorem} \overline{\text{``}\star\star\text{''}}]\text{''}} \supset \text{Theorem} \overline{\text{``}[\sim \star\star]\text{''}}$ 7108: a
- (d2) $b1 \vdash \text{Theorem} \overline{\text{``}[\sim \star\star]\text{''}}$ MP: c1, d1
- (e1) $b1 \vdash \sim \text{Consis}$ Rule P: 7109, b1, d2
- (f1) $\vdash \text{Theorem} \overline{\text{``}\star\star\text{''}} \supset \sim \text{Consis}$ 5240: e1
- (f2) $\vdash \text{Consis} \supset \sim \text{Theorem} \overline{\text{``}\star\star\text{''}}$ Rule P: f1
- (g) $\vdash \text{Consis} \supset \star\star$ Rule P: a, f2

Now suppose $\vdash \text{Consis}$. Then $\vdash \star\star$, so by (f), \mathcal{A} is inconsistent. Thus if \mathcal{A} is consistent, not $\vdash \text{Consis}$. ■

Let us consider again the proof of 7102. By a purely syntactic argument we were able to prove for every natural number k that $\vdash \sim \text{Proof}_{o\sigma\sigma} \text{``}\star\star\text{k''}$. Why are we unable to prove $\forall y_\sigma \sim \text{Proof}_{o\sigma\sigma} \text{``}\star\star\text{''}y$ (which is equivalent to $\star\star$)? Simply because we cannot prove *Consis* in the object language.

Gödel's Second (Incompleteness) Theorem shows that objective (2) of Hilbert's program, as originally conceived, is impossible to achieve for most mathematically significant theories. However, the significance of the theorem extends far beyond this. We next discuss, in an informal way, some additional consequences of the theorem. These considerations are quite general, and for purposes of this discussion we shall ignore various fine points, such as the differences between various formulations of type theory (such as the systems \mathcal{F}^ω and \mathcal{Q}_0 of §50, the systems \mathcal{T} and \mathcal{C} of X8030, and other systems in the literature) and certain technicalities regarding different orders of logic. (See [Andrews, 1974b] and [Gandy, 1956] for more details.) We shall temporarily use \mathcal{T}^∞ as a somewhat ambiguous name for a system obtained by adding an axiom of infinity such as Axiom 6 to a system of type theory.

1. It is very natural to extend the system \mathcal{G}^+ discussed in §31 to higher-order logic, and then to try to extend Gentzen's *Hauptsatz* to higher-order logic by showing that the Cut rule is not actually needed. The conjecture that this could be done was known as Takeuti's Conjecture because of Takeuti's extensive work related to it. Takeuti also pointed out in [Takeuti, 1953] and [Takeuti, 1958] that this Cut-Elimination conjecture implied, by a quite elementary argument, the consistency of \mathcal{T}^∞ . Therefore, by Gödel's Second Theorem one could not hope to prove the Cut-Elimination Theorem for type theory by methods which could be formalized in \mathcal{T}^∞ . Indeed, when proofs [Takahashi, 1967] [Prawitz, 1968] for the Cut-Elimination Theorem for type theory were found, they used semantic methods whose formalization would require a richer language than (finite) type theory, such as transfinite type theory or axiomatic set theory.
2. In automated theorem proving it is useful to work with a variety of formats for proofs and translate from one to another, since the best context for seeking a proof does not necessarily correspond to the form in which one wishes to display the proof. A useful proof-format for searching for proofs and for analyzing certain theoretical issues is

an *expansion proof* [Miller, 1987]. It is desirable to prove that every proof of a theorem of type theory in natural-deduction format can be translated to an expansion proof for that theorem. However, if we had such a proof which could be formalized in \mathcal{T}^∞ , then we could use it to prove the consistency of \mathcal{T}^∞ in \mathcal{T}^∞ , contradicting Gödel's Second Theorem. (See [Andrews, 1989] for more details.) Thus, the study of such translations is intrinsically difficult.

3. It is often appropriate to compare the strengths of logical systems in which mathematics can be formalized by saying that the system \mathcal{L}_2 is *stronger* than the system \mathcal{L}_1 if the consistency of \mathcal{L}_1 can be proved in \mathcal{L}_2 . This relation is generally found to be transitive, and (because of Gödel's Second Theorem) irreflexive (provided the systems under consideration are actually consistent). For example, Kemeny [Kemeny, 1949] showed that Zermelo Set Theory (which is weaker than the popular Zermelo-Fraenkel Set Theory) is stronger than \mathcal{T}^∞ by proving the consistency of \mathcal{T}^∞ in Zermelo Set Theory. This contradicted the widely held assumption that the two systems were of essentially the same strength. For another example, let us call *nth-order arithmetic* the result of adding an Axiom of Infinity (or Peano's Postulates) to n -th order logic. One can prove the consistency of n th-order arithmetic in $(n+2)$ th-order arithmetic, which establishes that $(n+2)$ th-order arithmetic is stronger than n th-order arithmetic. As was mentioned in the discussion following theorem 5402, Gödel asserted in [Gödel, 1936] that “passing to the logic of the next higher order has the effect, not only of making provable certain propositions that were not provable before, but also of making it possible to shorten, by an extraordinary amount, infinitely many of the proofs already available”. A complete proof of this may be found in [Buss, 1994].
4. In automated theorem proving, one tries to constrain as much as possible the space one must search while looking for a proof of a theorem. In particular, when seeking an appropriate wff of the form $[\lambda v_\beta A_o]$ with which to instantiate a quantifier such as $\forall u_{\alpha\beta}$, one would like to constrain the possible choices for A_o as much as possible. A_o may contain connectives and quantifiers, and it is natural to ask whether one can limit the orders of the variables in these new quantifiers to be no greater than the orders of the variables which occur in the theorem one is trying to prove. When the question is posed in this general a form, the answer is negative. For each positive integer n , let $Consis_o^n$ be a

sentence (similar to Consis_o^A) which expresses the consistency of n th-order arithmetic. Note that this is simply an assertion about numbers which can be expressed in some low-order logic which is independent of n . Since Consis_o^n is provable in $(n+2)$ th-order arithmetic, [$\text{Axiom6} \supset \text{Consis}_o^n$] is provable in $(n+2)$ th-order logic, so the wffs [$\text{Axiom6} \supset \text{Consis}_o^{10}$], [$\text{Axiom6} \supset \text{Consis}_o^{12}$], [$\text{Axiom6} \supset \text{Consis}_o^{14}$], ... are all theorems of type theory which are wffs of the same fixed low-order logic, but by Gödel's Second Theorem their proofs must be carried out in logics of increasingly high orders.

The proofs of Gödel's Incompleteness Theorems have been based upon the fact that it is possible to construct a sentence which can be interpreted as saying "I am not provable". It is also possible to construct a sentence which says "I am provable". Henkin [Henkin, 1952] raised the question whether this sentence is provable. This sentence does not seem to suggest any paradox, and at first glance it seems equally plausible that it is or is not provable. Nevertheless, Löb soon showed [Löb, 1955] that it is provable. We shall next present Löb's proof of this fact, and show how an alternative proof of Gödel's Second Theorem follows from Löb's Theorem.

We need a few more preliminary results, all of which follow from the following lemma:

7111 Lemma.

$$\vdash \dot{\forall}x_\sigma \dot{\forall}y_\sigma \blacksquare \text{Theorem}_{oo\sigma}[\overline{[[\supset_{ooo}]]} * x_\sigma * \overline{2048} * y_\sigma * \overline{2048}] \\ \supset \blacksquare \text{Theorem}_{oo\sigma}x_\sigma \supset \text{Theorem}_{oo\sigma}y_\sigma.$$

This lemma expresses within \mathcal{A} the fact that Modus Ponens (5224) is a derived rule of inference. The proof, which we omit, is similar in style to the proof of 7106, though longer.

7112 Corollary. For any wffs_o **A** and **B**, $\vdash \text{Theorem}_{oo\sigma}\overline{["A \supset B"]} \supset \blacksquare \text{Theorem}_{oo\sigma}\overline{"A"} \supset \text{Theorem}_{oo\sigma}\overline{"B"}$.

Proof: By 7111. ■

7113 Corollary. If $\vdash A \supset B$, then $\vdash \text{Theorem}_{oo\sigma}\overline{"A"} \supset \text{Theorem}_{oo\sigma}\overline{"B"}$.

Proof: By 7103 and 7112. ■

The answer to Henkin's question follows from the next theorem. Some insight into the idea motivating the proof of this theorem may be gained by

identifying the notion of provability with that of truth, and considering the following paradoxical argument in informal English: Given any statement \mathbf{D} , let \mathbf{E} be the statement “If I am true, so is \mathbf{D} .” Clearly if \mathbf{E} is true, then \mathbf{D} is true. Hence \mathbf{E} is true. Therefore \mathbf{D} is true, for *any* statement \mathbf{D} .

7114 Löb's Theorem. For any sentence \mathbf{D} , $\vdash \mathbf{D}$ if and only if $\vdash \text{Theorem}_{oo} \neg \overline{\mathbf{D}} \supset \mathbf{D}$.

Proof: To prove the theorem in the nontrivial direction, let \mathbf{G}_{oo} be the wff $[\lambda x_\sigma \neg \text{Theorem}_{oo} [128 * x_\sigma * [\text{Num}_{\sigma\sigma} x_\sigma] * 2048] \supset \mathbf{D}]$. Let $m = \langle \mathbf{G}_{oo} \rangle$ and let \mathbf{E}_o be $[\mathbf{G}_{oo} \overline{m}]$.

- (.1) $\vdash \mathbf{E} = \neg \text{Theorem} \neg \overline{\mathbf{E}} \supset \mathbf{D}$ λ , def of \mathbf{E} , representability of $*$ and Num
- (.2) $\vdash \text{Theorem} \neg \overline{\mathbf{E}} \supset \neg \text{Theorem} [\text{Theorem} \neg \overline{\mathbf{E}} \supset \mathbf{D}]$ 7108: .1
- (.3) $\vdash \text{Theorem} [\text{Theorem} \neg \overline{\mathbf{E}} \supset \mathbf{D}] \supset \neg \text{Theorem} [\text{Theorem} \neg \overline{\mathbf{E}}]$ $\supset \text{Theorem} \neg \overline{\mathbf{D}}$ 7112
- (.4) $\vdash \text{Theorem} \neg \overline{\mathbf{E}} \supset \text{Theorem} [\text{Theorem} \neg \overline{\mathbf{E}}]$ 7105
- (.5) $\vdash \text{Theorem} \neg \overline{\mathbf{E}} \supset \text{Theorem} \neg \overline{\mathbf{D}}$ Rule P: .2, .3, .4

Now suppose

- (.6) $\vdash \text{Theorem} \neg \overline{\mathbf{D}} \supset \mathbf{D}$

Then

- (.7) $\vdash \text{Theorem} \neg \overline{\mathbf{E}} \supset \mathbf{D}$ Rule P: .5, .6
- (.8) $\vdash \mathbf{E}$ Rule P: .1, .7
- (.9) $\vdash \text{Theorem} \neg \overline{\mathbf{E}}$ 7103: .8
- (.10) $\vdash \mathbf{D}$ Rule P: .7, .9

■

7115 Corollary. If \mathbf{D} is any sentence such that $\vdash \mathbf{D} = \text{Theorem} \neg \overline{\mathbf{D}}$, then $\vdash \mathbf{D}$.

Proof: By 7114. ■

Note that if \mathbf{D} is any such sentence, then \mathbf{D} asserts its own provability. Of course we can easily construct such a sentence. Simply drop the negation from the definition of $\star\star$ in the proof of 7110.

REMARK. We now have an elegant alternative proof of Gödel's Second Theorem. Suppose $\vdash \text{Consis}$, i.e., $\vdash F_o = \text{Theorem} \neg \overline{F_o}$ (by the definitions of Consis and \sim). Then $\vdash F_o$ by 7115, so \mathcal{A} is inconsistent.

Note that the proofs of both Gödel's Second Theorem (in either version which we have given) and Löb's Theorem can be given in full detail once one carries out the details of the proofs of 7104 and 7111. If one proves 7111, one can avoid using 7106.

EXERCISES

X7100. Show that if \mathcal{A} is ω -consistent and complete, then \mathcal{A} is ω -complete.

X7101. Show that if \mathcal{A} is consistent and ω -complete, then \mathcal{A} is ω -consistent.

X7102. Use the ideas in the proofs of 7101 and 7102 to construct a theory which is consistent (if \mathcal{Q}_0^∞ is consistent) but ω -inconsistent.

X7103. Let \mathcal{A} be an ω -consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Let $B_{o\sigma}$ be $[\lambda x_\sigma \bullet [x = \bar{0} \wedge \text{Consis}^{\mathcal{A}}] \vee \bullet x = \bar{1} \wedge \sim \text{Consis}^{\mathcal{A}}]$. Prove that $\vdash_{\mathcal{A}} \exists x_\sigma B_{o\sigma} x_\sigma$ but there is no natural number n such that $\vdash_{\mathcal{A}} B_{o\sigma} \bar{n}$.

X7104. Show that if \mathbf{D} is any wff_o, $\vdash \mathbf{D}$ iff $\vdash \text{Theorem}^{\overline{\mathbf{D}}} = \mathbf{D}$.

X7105. Show that for any cwff $\mathbf{D}_{o\sigma}$, there is a sentence \mathbf{E}_o such that $\vdash \mathbf{E}_o = \mathbf{D}_{o\sigma} \overline{\mathbf{E}_o}$.

X7106. Show that if \mathcal{A} is a consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ , there is no sentence \mathbf{B} such that $\vdash \sim \text{Theorem}^{\overline{\mathbf{B}}}$.

X7107. Prove Gödel's Second Theorem by constructing a sentence \mathbf{E} such that $\vdash \mathbf{E} = \text{Theorem}^{\overline{\sim \mathbf{E}}}$. (*Hint:* prove $\vdash \mathbf{E} \supset \bullet \text{Theorem}^{\overline{\mathbf{E}}} \wedge \text{Theorem}^{\overline{\sim \mathbf{E}}}$.)

X7108. Let \mathcal{A} be a consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ . If $\vdash_{\mathcal{A}} \mathbf{E} = \sim \text{Theorem}_{o\sigma}^{\mathcal{A}} \overline{\sim \mathbf{E}}$, then \mathbf{E} expresses the consistency of the system obtained by adding \mathbf{E} to \mathcal{A} as an additional axiom. Is \mathbf{E} consistent with \mathcal{A} ?

§72. Essential Incompleteness

We start by showing that no extension of \mathcal{Q}_0^∞ can be consistent, recursively axiomatized, and complete.

7200 Gödel-Rosser Incompleteness Theorem. [Rosser, 1936] Let \mathcal{A} be a consistent, recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Let $\nabla_{o\sigma}$ be

$$[\lambda x_\sigma \dot{\forall} y_\sigma \bullet \text{Proof}_{o\sigma\sigma}^{\mathcal{A}}[\overline{128} * x_\sigma * [\text{Num}_{\sigma\sigma} x_\sigma] * \overline{2048}]y_\sigma \supset$$

$$\dot{\exists} z_\sigma \bullet z_\sigma \leq y_\sigma \wedge \text{Proof}_{o\sigma\sigma}^{\mathcal{A}}[\overline{["\sim_{oo} [" * x_\sigma * [\text{Num}_{\sigma\sigma} x_\sigma] * "]]}]z_\sigma.$$

Let $m = "\nabla_{o\sigma}"$ and Δ_o be $[\nabla_{o\sigma} \overline{m}]$. Neither $\vdash_{\mathcal{A}} \Delta_o$ nor $\vdash_{\mathcal{A}} \sim \Delta_o$.

Proof:

$$(.) \vdash [\overline{128} * \overline{m} * [\text{Num } \overline{m}] * \overline{2048}] = \overline{\Delta}$$

since $*_{\sigma\sigma\sigma}$ and $\text{Num}_{\sigma\sigma}$ represent the corresponding functions.

$$(.) \vdash \overline{["\sim_{oo} [" * \overline{m} * [\text{Num } \overline{m}] * "]]} = \overline{["\sim \Delta"]}$$

for similar reasons.

$$(.) \vdash \Delta = \dot{\forall} y_\sigma \bullet \text{Proof}^{\overline{\Delta}}y_\sigma \supset \dot{\exists} z_\sigma \bullet z_\sigma \leq y_\sigma \wedge \text{Proof}^{\overline{["\sim \Delta"]}}z_\sigma$$

λ , def of Δ ; R: .1, .2

Suppose (.) $\vdash \Delta$. Let n be the number of a proof of Δ .

$$(.) \vdash \text{Proof}^{\overline{\Delta}}\overline{n} \quad \text{since } \text{Proof}_{o\sigma\sigma} \text{ represents the relation Proof.}$$

$$(.) \vdash \dot{\forall} y_\sigma \bullet \text{Proof}^{\overline{\Delta}}y_\sigma \supset \dot{\exists} z_\sigma \bullet z_\sigma \leq y_\sigma \wedge \text{Proof}^{\overline{["\sim \Delta"]}}z_\sigma \quad \text{R: .4, .3}$$

$$(.) \vdash \text{Proof}^{\overline{\Delta}}\overline{n} \supset \dot{\exists} z_\sigma \bullet z_\sigma \leq \overline{n} \wedge \text{Proof}^{\overline{["\sim \Delta"]}}z_\sigma \quad \text{VI: .6; 6110}$$

$$(.) \vdash \dot{\exists} z_\sigma \bullet z \leq \overline{n} \wedge \text{Proof}^{\overline{["\sim \Delta"]}}z_\sigma \quad \text{MP: .5, .7}$$

$$(.) \vdash \text{Proof}^{\overline{["\sim \Delta"]}}0 \vee \dots \vee \text{Proof}^{\overline{["\sim \Delta"]}}\overline{n} \quad \text{Sub: 6210}^n; \text{R: .8}$$

Since \mathcal{A} is consistent, by .4 we have not $\vdash \sim \Delta$. Hence for each natural number k , k is not the number of a proof of $\sim \Delta$, so for each natural number k ,

$$(.) \vdash \sim \text{Proof}^{\overline{["\sim \Delta"]}}\overline{k}$$

$$(.) \vdash F_o \quad \text{Rule P: .9, .10}^0 - .10^n$$

This contradicts the consistency of \mathcal{A} ; hence Δ cannot be a theorem.

Next suppose (.) $\vdash \sim \Delta$. Let n be the number of a proof of $\sim \Delta$.

$$(.) \vdash \text{Proof}^{\overline{["\sim \Delta"]}}\overline{n} \quad \text{by representability}$$

$$(.) \vdash \dot{\exists} y_\sigma \bullet \text{Proof}^{\overline{\Delta}}y_\sigma \wedge \dot{\forall} z_\sigma \sim \bullet z_\sigma \leq y_\sigma \wedge \text{Proof}^{\overline{["\sim \Delta"]}}z_\sigma$$

Rule Q: .12, .3

$$(.) \vdash \dot{\exists} y_\sigma \bullet \text{Proof}^{\overline{\Delta}}y_\sigma \wedge \dot{\forall} z_\sigma \sim \bullet z_\sigma \leq y_\sigma \wedge \text{Proof}^{\overline{["\sim \Delta"]}}z_\sigma$$

Choose y

- (.16) $.15 \vdash \sim \cdot \bar{n} \leq y_\sigma \wedge \text{Proof}^{\overline{[\sim \Delta]}} \bar{n}$ Rule P, VI: .15, 6110
(.17) $.15 \vdash \bar{n} \leq y_\sigma \vee y_\sigma \leq \bar{n}$ VI: 6212; Rule P: 6110, .15
(.18) $.15 \vdash \mathbb{N} y_\sigma \wedge y_\sigma \leq \bar{n} \wedge \text{Proof}^{\overline{\Delta}} y_\sigma$ Rule P: .13, .15, .16, .17
(.19) $.15 \vdash \dot{\exists} y_\sigma \cdot y_\sigma \leq \bar{n} \wedge \text{Proof}^{\overline{\Delta}} y_\sigma$ \exists Gen: .18
(.20) $\vdash .19$ Rule C: .14, .19
(.21) $\vdash \text{Proof}^{\overline{\Delta}} 0 \vee \dots \vee \text{Proof}^{\overline{\Delta}} \bar{n}$ Sub: 6210ⁿ; R: .20

Since \mathcal{A} is consistent, by .12 we cannot have $\vdash \Delta$, so for each natural number k ,

- (.22^k) $\vdash \sim \text{Proof}^{\overline{\Delta}} k$
(.23) $\vdash F_o$ Rule P: .21, .22⁰ – .22ⁿ

This contradicts the consistency of \mathcal{A} , so $\sim \Delta$ cannot be a theorem. ■

Note that with respect to our Gödel numbering, Δ says “If I am provable, so is my negation.” Whether or not \mathcal{A} is consistent, this is true. It may help to understand this proof if we observe that $\sim \Delta$ also asserts (or at least implies) “If I am provable, so is my negation.” We can see this as follows:

- (.24) $\vdash \sim \Delta \equiv \dot{\exists} y_\sigma \cdot \text{Proof}^{\overline{\Delta}} y_\sigma \wedge \dot{\forall} z_\sigma \cdot \sim z_\sigma \leq y_\sigma \vee \sim \text{Proof}^{\overline{[\sim \Delta]}} z_\sigma$ Rule Q: .3
(.25) $\vdash \sim \Delta \equiv \dot{\exists} y_\sigma \dot{\forall} z_\sigma \cdot \text{Proof}^{\overline{\Delta}} y_\sigma \wedge \sim z_\sigma \leq y_\sigma \vee \sim \text{Proof}^{\overline{[\sim \Delta]}} z_\sigma$ Rule Q (prenex): .24
(.26) $\vdash \sim \Delta \supset \dot{\forall} z_\sigma \dot{\exists} y_\sigma \cdot \text{Proof}^{\overline{\Delta}} y_\sigma \wedge \sim z_\sigma \leq y_\sigma \vee \sim \text{Proof}^{\overline{[\sim \Delta]}} z_\sigma$ Rule Q (2137): .25
(.27) $\vdash \sim \Delta \supset \dot{\forall} z_\sigma \dot{\exists} y_\sigma \cdot \sim \text{Proof}^{\overline{[\sim \Delta]}} z_\sigma \vee \sim z_\sigma \leq y_\sigma \wedge \text{Proof}^{\overline{\Delta}} y_\sigma$ Rule Q (Rule P, 2105): .26
(.28) $\vdash \sim \Delta \supset \dot{\forall} z_\sigma \dot{\exists} y_\sigma \cdot \text{Proof}^{\overline{[\sim \Delta]}} z_\sigma \supset \sim y_\sigma \leq z_\sigma \wedge \text{Proof}^{\overline{\Delta}} y_\sigma$ Rule Q: .27, 6212
(.29) $\vdash \sim \Delta \supset \dot{\forall} z_\sigma \cdot \text{Proof}^{\overline{[\sim \Delta]}} z_\sigma \supset \dot{\exists} y_\sigma \cdot y_\sigma \leq z_\sigma \wedge \text{Proof}^{\overline{\Delta}} y_\sigma$ Rule Q (antiprenex): .28

Note how (.29) mirrors (.3). Of course, (.3) is an equivalence and (.29) is only an implication, but the implication suffices for proving that $\sim \Delta$ is not a theorem just as it was proved that Δ is not a theorem.

REMARK. As noted after the proof of theorem 7009, the requirement that \mathcal{A} be a *pure* extension of \mathcal{Q}_0^∞ can easily be dropped. Thus we obtain the following:

7201 Corollary.

- (a) If \mathcal{A} is any consistent and complete extension of \mathcal{Q}_0^∞ , then \mathcal{A} is not recursively axiomatizable.
- (b) If \mathcal{M} is any model of \mathcal{Q}_0^∞ , the set of sentences true in \mathcal{M} is not recursively axiomatizable.
- (c) If \mathcal{M} is any model of \mathcal{Q}_0^∞ , the set of sentences true in \mathcal{M} is not recursive.

It is well known in informal mathematics that Peano's Postulates are categorical. In our present context, this means that all *standard* models of Peano's Postulates are isomorphic. We shall sharpen this statement slightly, and then use it with 7200 to show, as promised at the end of §55, that there is no recursively axiomatized pure extension of \mathcal{Q}_0 in which the theorems are precisely the wffs_o which are valid in the standard sense. The sentence Inf^{III} below states Peano's Postulates and fixes the behavior of the description operator in an arbitrary but conveniently simple manner. (Recall that the wffs $\Pi_{o(o)}$ and $\sum_{o(o)}^1$ used below were defined in §51 and §53, respectively.)

DEFINITION. Let Inf^{III} be the sentence

$$\begin{aligned} \exists o_i \exists s_{ii} [\forall x_i [s_{ii}x_i \neq o_i] \wedge \forall x_i \forall y_i [s_{ii}x_i = s_{ii}y_i \supset x_i = y_i] \\ \wedge \forall p_{oi} [[p_{oi}o_i \wedge \forall x_i \cdot p_{oi}x_i \supset p_{oi} \cdot s_{ii}x_i] \supset \Pi_{o(o)}p_{oi}] \\ \wedge \forall p_{oi} \cdot \sim \sum_{o(o)}^1 p_{oi} \supset \iota_{i(o)}p_{oi} = o_i]. \end{aligned}$$

7202 Proposition. $\vdash_{\mathcal{Q}_0} Inf^{III} \supset$ Axiom 6.

Proof: It is easy to see that $\vdash_{\mathcal{Q}_0} Inf^{III} \supset Inf^{II}$ for the wff Inf^{II} of §60. Simply let n_{oi} be $[\lambda x_i T_o]$ in Inf^{II} . Since (as discussed in §60 and §62), $\vdash_{\mathcal{Q}_0} Inf^{II} \supset$ Axiom 6, the result follows easily. (For an alternative approach, see X7200.) ■

We next show that Inf^{III} has an essentially unique standard model. Recall that the concept of isomorphism of models was defined in §54.

7203 Proposition. There is a standard model \mathcal{N} for $\{\text{Inf}^{III}\}$ such that \mathcal{N} is isomorphic to every standard model for $\{\text{Inf}^{III}\}$.

Proof: Let \mathcal{E}_i be the set \mathbb{N} of natural numbers. Let $\mathcal{N} = \langle \{\mathcal{E}_\alpha\}_\alpha, \mathcal{J} \rangle$ be the standard model with \mathcal{E}_i as domain of individuals, and with $\mathcal{J}_{\iota_i(o_i)}$ as that function from \mathcal{E}_o into \mathcal{E}_i which maps each one-element set to its unique member, and every other set to the number zero.

Using X5403 it is straightforward to show, in the manner of the proof of 5402 (the Soundness Theorem), that $\mathcal{N} \models \text{Inf}^{III}$. We leave the detailed verification to the reader. Thus \mathcal{N} is a model for $\{\text{Inf}^{III}\}$.

Next suppose $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{I} \rangle$ is any standard model for Inf^{III} . Thus there is an assignment φ into \mathcal{M} such that

$$\begin{aligned} \mathcal{M} \models_\varphi [\forall x_i [s_{ii}x_i \neq o_i] \wedge \forall x_i \forall y_i [s_{ii}x_i = s_{ii}y_i \supset x_i = y_i] \\ \wedge \forall p_{oi} [[p_{oi}o_i \wedge \forall x_i \bullet p_{oi}x_i \supset p_{oi} \bullet s_{ii}x_i] \supset \Pi_{o(oi)}p_{oi}] \\ \wedge \forall p_{oi} \bullet \sim \sum_{o(oi)}^1 p_{oi} \supset \iota_{i(oi)}p_{oi} = o_i]. \end{aligned}$$

We define a map τ from \mathcal{E}_i to \mathcal{D}_i by recursion: $\tau 0 = \varphi o_i$, and $\tau(Sn) = (\varphi s_n)(\tau n)$ for each $n \in \mathbb{N}$.

We prove τ is one-one by contradiction. Suppose τ is not one-one, and k is the least natural number such that there exists $j \in \mathbb{N}$ such that $j \neq k$ and $\tau j = \tau k$. Clearly $k < j$, so there exists $m \in \mathbb{N}$ such that $j = Sm$. $\mathcal{M} \models_\varphi \forall x_i [s_{ii}x_i \neq o_i]$, so $\mathcal{M} \models_{(\varphi : x_i / \tau m) \sim} [s_{ii}x_i = o_i]$, so $\tau k = \tau j = \tau(Sm) = (\varphi s_n)(\tau m) \neq \varphi o_i = \tau 0$, so $k \neq 0$. Hence there exists $n \in \mathbb{N}$ such that $k = Sn$. Let $\psi = ((\varphi : x_i / \tau m) : y_i / \tau n)$. $\mathcal{V}_\psi[s_{ii}x_i] = (\varphi s_n)(\tau m) = \tau(Sm) = \tau j = \tau k = \tau(Sn) = (\varphi s_n)(\tau n) = \mathcal{V}_\psi[s_{ii}y_i]$, so $\mathcal{M} \models_\psi [s_{ii}x_i = s_{ii}y_i]$. However, $\mathcal{M} \models_\varphi \forall x_i \forall y_i [s_{ii}x_i = s_{ii}y_i \supset x_i = y_i]$, so $\mathcal{M} \models_\psi [s_{ii}x_i = s_{ii}y_i \supset x_i = y_i]$, so $\mathcal{M} \models_\psi [x_i = y_i]$, so $\tau m = \psi x_i = \psi y_i = \tau n$. Since $n < Sn = k$, this contradicts the choice of k . Thus τ must be one-one on \mathcal{E}_i .

To show that τ maps \mathcal{E}_i onto \mathcal{D}_i , let u be that function from \mathcal{D}_i into \mathcal{D}_o such that for each $d \in \mathcal{D}_i$, $ud = \top$ iff d is in the range of τ . (Informally, we may regard u as the range of τ). Since \mathcal{M} is a standard model, $u \in \mathcal{D}_o$. Let $\psi = (\varphi : p_{oi}/u)$. $\mathcal{M} \models_\psi [[p_{oi}o_i \wedge \forall x_i \bullet p_{oi}x_i \supset p_{oi} \bullet s_{ii}x_i] \supset \Pi_{o(oi)}p_{oi}]$. $\tau 0 = \varphi o_i$, so $\mathcal{V}_\psi^M[p_{oi}o_i] = u(\varphi o_i) = \top$. Also, for any $d \in \mathcal{D}_i$, if $ud = \top$, then there exists $n \in \mathcal{E}_i$ such that $\tau n = d$, so $\tau(Sn) = (\varphi s_n)(\tau n) = (\varphi s_n)d$, so $u((\varphi s_n)d) = \top$; thus $\mathcal{M} \models_\psi \forall x_i [p_{oi}x_i \supset p_{oi} \bullet s_{ii}x_i]$. Therefore $\mathcal{M} \models_\psi \Pi_{o(oi)}p_{oi}$, i.e., $\mathcal{V}_\psi[\lambda x_i T_o] = u$. Thus u maps every element of \mathcal{D}_i to \top , so \mathcal{D}_i is the range of τ .

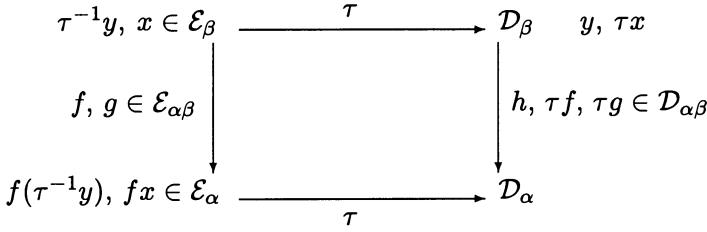


Figure 7.1: Diagram for proof of 7203

By induction on α , for each type symbol α we next extend τ to map \mathcal{E}_α into \mathcal{D}_α , and show that τ is a one-one map from \mathcal{E}_α onto \mathcal{D}_α .

Since $\mathcal{E}_o = \{\top, \perp\} = \mathcal{D}_o$, we let τ be the identity map on \mathcal{E}_o .

Given that τ is a one-one map from \mathcal{E}_α onto \mathcal{D}_α , and from \mathcal{E}_β onto \mathcal{D}_β , we define τ from $\mathcal{E}_{\alpha\beta}$ into $\mathcal{D}_{\alpha\beta}$ as follows: given $f \in \mathcal{E}_{\alpha\beta}$, let τf be that function from \mathcal{D}_β to \mathcal{D}_α such that for each $y \in \mathcal{D}_\beta$, $(\tau f)y = \tau(f(\tau^{-1}y))$. (See Figure 7.1). Thus for any $x \in \mathcal{E}_\beta$, $(\tau f)(\tau x) = \tau(f(\tau^{-1}\tau x)) = \tau(fx)$. Clearly $\tau f \in \mathcal{D}_{\alpha\beta}$, since \mathcal{M} is a standard model. To show τ is one-one on $\mathcal{E}_{\alpha\beta}$, suppose f and g are in $\mathcal{E}_{\alpha\beta}$ and $\tau f = \tau g$. For any $x \in \mathcal{E}_\beta$, $\tau(fx) = (\tau f)(\tau x) = (\tau g)(\tau x) = \tau(gx)$, so $fx = gx$ since τ is one-one on \mathcal{E}_α ; hence $f = g$.

To show that τ maps $\mathcal{E}_{\alpha\beta}$ onto $\mathcal{D}_{\alpha\beta}$, let h be any member of $\mathcal{D}_{\alpha\beta}$. Define f from \mathcal{E}_β to \mathcal{E}_α so that for each $x \in \mathcal{E}_\beta$, $fx = \tau^{-1}(h(\tau x))$. $f \in \mathcal{E}_{\alpha\beta}$ since \mathcal{N} is standard. Then for any $y \in \mathcal{D}_\beta$, $(\tau f)y = \tau(f(\tau^{-1}y)) = \tau\tau^{-1}h(\tau\tau^{-1}y) = hy$, so $\tau f = h$.

It remains only to show that for each primitive constant c_α of \mathcal{Q}_0 , $\tau(\mathcal{J}c_\alpha) = \mathcal{I}c_\alpha$. Suppose $c_\alpha = Q_{\alpha\beta\beta}$; then $\mathcal{J}c_\alpha$ is the identity relation q on \mathcal{E}_β ; we must show that τq is the identity relation $\mathcal{I}Q_{\alpha\beta\beta}$ on \mathcal{D}_β . For any z and y in \mathcal{D}_β , $((\tau q)z)y = (\tau(q(\tau^{-1}z)))y = \tau((q(\tau^{-1}z))(\tau^{-1}y)) = q(\tau^{-1}z)(\tau^{-1}y) = \top$ iff $\tau^{-1}z = \tau^{-1}y$ iff $z = y$, so τq is the identity relation on \mathcal{D}_β .

Finally we show that $\tau(\mathcal{J}\iota_{\iota(o_i)}) = \mathcal{I}\iota_{\iota(o_i)}$. Let $u \in \mathcal{D}_{o_i}$. Since τ is the identity map on \mathcal{E}_o , for each $y \in \mathcal{D}_i$, $(\tau^{-1}u)(\tau^{-1}y) = \tau((\tau^{-1}u)(\tau^{-1}y)) = (\tau\tau^{-1}u)(\tau\tau^{-1}y) = uy$, so since τ is a one-one correspondence, $\tau^{-1}u$ is a one-element set iff u is, and if $u = \{y\}$, then $\tau^{-1}u = \{\tau^{-1}y\}$. Hence if $u = \{y\}$, $(\tau\mathcal{J}\iota_{\iota(o_i)})u = \tau((\mathcal{J}\iota_{\iota(o_i)})(\tau^{-1}u)) = \tau(\tau^{-1}y) = y = (\mathcal{I}\iota_{\iota(o_i)})u$. Suppose u is not a one-element set; then $\tau^{-1}u$ is not either, so $(\tau\mathcal{J}\iota_{\iota(o_i)})u = \tau((\mathcal{J}\iota_{\iota(o_i)})(\tau^{-1}u)) = \tau 0 = \varphi_{o_i} = (\mathcal{I}\iota_{\iota(o_i)})u$ since $\mathcal{M} \models_{\varphi} \forall p_{oi} \cdot \sim \sum_{o(o_i)}^1 p_{oi} \supset \iota_{\iota(o_i)} p_{oi} = o_i$. Thus $\tau(\mathcal{J}\iota_{\iota(o_i)}) = \mathcal{I}\iota_{\iota(o_i)}$. ■

7204 Corollary. \mathcal{Q}_0^∞ is consistent.

Proof: By 7203 $\mathcal{N} \models Inf^{III}$, but by 7202 and 5402, $\mathcal{N} \models Inf^{III} \supset Axiom\ 6$. Hence $\mathcal{N} \models Axiom\ 6$. Therefore \mathcal{N} is a model of \mathcal{Q}_0^∞ , and \mathcal{Q}_0^∞ is consistent by 5403. ■

Naturally, the consistency proof above is a semantic one, which can be carried out only in a language which is stronger than \mathcal{Q}_0^∞ , and hence more likely to be inconsistent. Therefore it cannot bring much reassurance to anyone who has serious doubts about the existence of the set of natural numbers or the possibility of reasoning in a consistent way about sets and functions. However, it may be reassuring to one who has no such fundamental doubts, but who has been concerned about the possibility of serious flaws in our formalization of these notions in \mathcal{Q}_0^∞ .

DEFINITION. If \mathcal{A} be any extension of \mathcal{Q}_0 , we say that \mathcal{A} is *essentially incomplete* iff \mathcal{A} is consistent and recursively axiomatized and every consistent recursively axiomatized pure extension of \mathcal{A} is incomplete.

REMARK. We include the requirement that \mathcal{A} be consistent and recursively axiomatized in the definition of essential incompleteness so that systems (such as inconsistent systems) which have no consistent recursively axiomatized pure extensions will not be regarded as essentially incomplete.

7205. Theorem. \mathcal{Q}_0^∞ is essentially incomplete.

Proof: By 7204 and 7200. ■

It can be argued that Axiom 6 is valid in every standard infinite model of \mathcal{Q}_0 , for it seems highly plausible that if \mathcal{M} is any standard model of \mathcal{Q}_0 , then the Axiom of Choice is true in \mathcal{M} , so there is a well-ordering of the individuals of \mathcal{M} in \mathcal{D}_{on} . If \mathcal{M} is infinite and standard, from this well-ordering one can construct a relation r satisfying the requirements of Axiom 6. Thus if one is interested primarily in standard models and one wishes a system with an axiom of infinity, Axiom 6 should be derivable, so by 7205 one must have a system which is inconsistent or incomplete.

REMARK. We saw in 5404 that every finite model of \mathcal{Q}_0 is a standard model. Also, for any positive integer n , let C^n be $\exists x_1^1 \dots \exists x_n^n [\bigvee_{1 \leq i \leq n} [y_i = x_i^i]] \wedge \bigwedge_{1 \leq i < j \leq n} [x^i \neq x^j] \wedge \forall p_{oi} [\sim \sum_{o(i)}^1 p_{oi} \supset \iota_{i(oi)}[\lambda x_i F_o] = \iota_{i(oi)} p_{oi}]$. Let \mathcal{Q}_0^n be the system obtained from \mathcal{Q}_0 by adding C^n as an axiom. It can be shown

that \mathcal{Q}_0^n is complete and consistent, and that all models of \mathcal{Q}_0^n are standard and isomorphic to one another. See Exercise X8032.

We now come to one of the most important theorems in this chapter.

7206 Theorem. There is no recursively axiomatized pure extension \mathcal{B} of \mathcal{Q}_0 such that the theorems of \mathcal{B} are precisely the wffs_o which are valid in the standard sense.

Proof: Suppose there is such a system \mathcal{B} . Let \mathcal{A} be the system obtained from \mathcal{B} by adding Inf^{III} as an additional axiom. Then $\text{Axiom}^{\mathcal{A}} = (\lambda n)[\text{Axiom}^{\mathcal{B}}(n) \vee n = \text{"}\text{Inf}^{III}\text{"}]$, and $\vdash_{\mathcal{A}} \text{Axiom 6}$ by 7202, so \mathcal{A} is a recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Since every standard model of \mathcal{Q}_0 is a model of \mathcal{B} , the standard model \mathcal{N} of 7203 is a standard model of \mathcal{A} , so \mathcal{A} is consistent (by 5403). Hence by 7205 \mathcal{A} is incomplete.

We next note that for any sentence \mathbf{C} , the following are equivalent:

- (a) $\vdash_{\mathcal{A}} \mathbf{C}$;
- (b) $\vdash_{\mathcal{B}} [\text{Inf}^{III} \supset \mathbf{C}]$;
- (c) $\mathcal{M} \models [\text{Inf}^{III} \supset \mathbf{C}]$ for every standard model \mathcal{M} ;
- (d) $\mathcal{N} \models \mathbf{C}$.

It is trivial that (a) \Leftrightarrow (b) and (b) \Leftrightarrow (c) and that (c) \Rightarrow (d) (since $\mathcal{N} \models \text{Inf}^{III}$). To show that (d) \Rightarrow (c), suppose that $\mathcal{N} \models \mathbf{C}$, and let \mathcal{M} be any standard model of \mathcal{Q}_0 . If $\mathcal{M} \models \text{Inf}^{III}$, then \mathcal{N} is isomorphic to \mathcal{M} by 7203, so $\mathcal{M} \models \mathbf{C}$ by 5407, so $\mathcal{M} \models \text{Inf}^{III} \supset \mathbf{C}$. On the other hand, if not $\mathcal{M} \models \text{Inf}^{III}$, then $\mathcal{M} \models \text{Inf}^{III} \supset \mathbf{C}$, so (c) follows.

Since for any sentence \mathbf{C} , either $\mathcal{N} \models \mathbf{C}$ or $\mathcal{N} \models \sim \mathbf{C}$, it follows that for any sentence \mathbf{C} , either $\vdash_{\mathcal{A}} \mathbf{C}$ or $\vdash_{\mathcal{A}} \sim \mathbf{C}$, so \mathcal{A} is complete. This contradiction shows that there is no such system \mathcal{B} . ■

7207 Corollary. There is no decision procedure for determining of an arbitrary sentence whether it is valid in the standard sense.

Proof: Suppose there were such a procedure. We could let \mathcal{B} be the extension of \mathcal{Q}_0 with postulates all those sentences which are valid in the standard sense. \mathcal{B} would be recursively axiomatized, so it would be a counterexample to Theorem 7206.

REMARK. We can now show that the Strong Compactness conjecture mentioned after Theorem 5503 is indeed false. We shall show that there is a set

\mathcal{S} of sentences such that every finite subset of \mathcal{S} has a standard model, but \mathcal{S} has no standard model.

Let \mathcal{Q}'_0 be the expansion of \mathcal{Q}_0 obtained by adding the constants \bar{o}_i and \bar{s}_n to the constants of \mathcal{Q}_0 , and let \mathcal{Q}''_0 be the expansion of \mathcal{Q}'_0 obtained by adding the constant ∞_i to the constants of \mathcal{Q}'_0 . Let $Inf^{III'}$ be the sentence

$$\begin{aligned} & [\forall x_i[\bar{s}_n x_i \neq \bar{o}_i] \wedge \forall x_i \forall y_i[\bar{s}_n x_i = \bar{s}_n y_i \supset x_i = y_i]] \\ & \wedge \forall p_{oi}[[p_{oi} \bar{o}_i \wedge \forall x_i p_{oi} x_i \supset p_{oi} \bar{s}_n x_i] \supset \Pi_{o(oi)} p_{oi}] \\ & \wedge \forall p_{oi} \sim \sum_{o(oi)}^1 p_{oi} \supset \iota_{i(oi)} p_{oi} = \bar{o}_i]. \end{aligned}$$

Let $\mathcal{S} = \{Inf^{III'}\} \cup \{\sim \infty_i = \bar{s}_n^n \bar{o}_i \mid n \text{ is a natural number}\}$

$$\begin{aligned} & = \{Inf^{III'}\} \cup \{\sim \infty_i = \bar{o}_i, \sim \infty_i = \bar{s}_n \bar{o}, \sim \infty_i = \bar{s}_n \bar{s}_n \bar{o}, \\ & \quad \sim \infty_i = \bar{s}_n \bar{s}_n \bar{s}_n \bar{o}, \dots\} \end{aligned}$$

Basically, an attempt to construct a standard model for \mathcal{S} fails because one cannot find a denotation for ∞_i , but the details of the argument must be handled carefully.

Let $\mathcal{N}' = \langle \{\mathcal{E}_\alpha\}_\alpha, \mathcal{J}' \rangle$ be the standard model (with \mathcal{Q}'_0 as language) obtained by extending the model \mathcal{N} in the proof of 7203 as follows: \mathcal{J}' agrees with \mathcal{J} on the constants of \mathcal{Q}_0 , $\mathcal{J}' \bar{o}_i = 0$, and $\mathcal{J}' \bar{s}_n$ is the successor function S on the natural numbers. Trivial modifications in the proof of 7203 establish that \mathcal{N}' is a standard model for $\{Inf^{III'}\}$ which is isomorphic to every standard model for $\{Inf^{III'}\}$. It is easily seen by induction that $\mathcal{V}^{\mathcal{N}'} \bar{s}_n^n \bar{o}_i = n$ for every natural number n .

To see that every finite subset of \mathcal{S} has a standard model, let \mathcal{S}_0 be any finite subset of \mathcal{S} . Define $\mathcal{N}'' = \langle \{\mathcal{E}_\alpha\}_\alpha, \mathcal{J}'' \rangle$, where \mathcal{J}'' is an extension of \mathcal{J}' obtained by defining $\mathcal{J}'' \infty_i$ to be a sufficiently large natural number that all of the (finitely many) sentences in \mathcal{S}_0 are true. Clearly \mathcal{N}'' is a standard model (with \mathcal{Q}''_0 as language) for \mathcal{S}_0 .

Now suppose \mathcal{S} has a standard model $\mathcal{M}'' = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{I}'' \rangle$ (which necessarily has \mathcal{Q}''_0 as language). Let \mathcal{I}' be the restriction of \mathcal{I}'' to the constants of \mathcal{Q}'_0 (so $\mathcal{I}' \infty_i$ is not defined), and let $\mathcal{M}' = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{I}' \rangle$. Then \mathcal{M}' is a standard model (with language \mathcal{Q}'_0) for $\{Inf^{III'}\}$, so there is an isomorphism τ from \mathcal{M}' to \mathcal{N}' .

$\mathcal{V}^{\mathcal{M}''} \infty_i \in \mathcal{D}_i$, so $\tau \mathcal{V}^{\mathcal{M}''} \infty_i$ is defined and is some natural number n in \mathcal{E}_i . $\tau \mathcal{V}^{\mathcal{M}''} \infty_i = n = \mathcal{V}^{\mathcal{N}'} \bar{s}_n^n \bar{o}_i = \tau \mathcal{V}^{\mathcal{M}'} \bar{s}_n^n \bar{o}_i$ (by 5406). Since τ is an injection,

$\mathcal{V}^{\mathcal{M}''}\infty_i = \mathcal{V}^{\mathcal{M}'}\bar{s}_n^n\bar{o}_i = \mathcal{V}^{\mathcal{M}''}\bar{s}_n^n\bar{o}_i$, so $\mathcal{M}'' \models \infty_i = \bar{s}_n^n\bar{o}_i$ by 5401(b). This contradicts the assumption that \mathcal{M}'' is a model for \mathcal{S} . Therefore, \mathcal{S} has no standard model.

EXERCISES

X7200. Give a direct proof of Proposition 7202 in the following way. Given o_i and s_n as in Inf^{III} , let R_{on} be $[\lambda x, \lambda y, \forall p_{on} \cdot [p_{on}[s_n x_i] \wedge \forall z_i \cdot p_{on} z_i \supset p_{on} s_n z_i] \supset p_{on} y_i]$. Show R_{on} has the properties required by Axiom 6. (Show $\sim R_{on} x_i x_i$ by induction on x_i .)

X7201. Show that every complete recursively axiomatized pure extension of \mathcal{Q}_0 is decidable (i.e., has a recursive set of theorems).

X7202. Show that there is a consistent recursively axiomatized pure extension \mathcal{Q}_0^∞ which has no standard models.

X7203. Is there a consistent recursively axiomatized extension of \mathcal{Q}_0 whose theorems include all sentences valid in all standard models?

X7204. Suppose \mathbf{A}_σ is a closed wff such that $\models N_{o\sigma} \mathbf{A}_\sigma$. Must there exist a natural number n such that $\vdash \mathbf{A}_\sigma = \bar{n}$?

§73. Undecidability and Undefinability

A logistic system is *decidable* iff there is an effective procedure for determining of an arbitrary wff of that system whether that wff is a theorem. When an effective Gödel numbering is given, the existence of such a procedure is equivalent to the existence of an effective procedure for determining of an arbitrary natural number whether it is the number of a theorem. By Church's Thesis, we can replace the notion of an effective procedure involving natural numbers by the notion of recursiveness. Thus we are led to the following definition:

DEFINITION. A pure extension \mathcal{A} of \mathcal{Q}_0 is (*recursively*) *decidable* iff $\{\text{"B"} \mid \vdash_{\mathcal{A}} \mathbf{B}\}$ is a recursive set of numbers.

7300 Theorem. If \mathcal{A} is any consistent pure extension of \mathcal{Q}_0^∞ , then \mathcal{A} is undecidable.

Proof: Suppose \mathcal{A} is decidable. Then the set \mathcal{S} of numbers of theorems of \mathcal{A} is recursive, and hence by 6401, there is a wff $Thm_{o\sigma}$ which represents this set. Let $G_{o\sigma}$ be $[\lambda x_\sigma \sim Thm_{o\sigma} * \overline{128} * x_\sigma * [Num_{\sigma\sigma} x_\sigma] * \overline{2048}]$. Let $m = "G_{o\sigma}"$ and let E_o be $[G_{o\sigma} \overline{m}]$. We shall write $\vdash \mathbf{B}$ for $\vdash_{\mathcal{A}} \mathbf{B}$.

$$(.) \vdash E = \sim Thm^{\overline{"E"}} \quad \text{def of } E$$

Suppose (.) $\vdash E$. Then

$$\begin{aligned} (.) &\vdash Thm^{\overline{"E"}} && \text{Since } "E" \in \mathcal{S} \\ (4) &\vdash \sim E && .1, .3 \end{aligned}$$

This contradicts the consistency of \mathcal{A} , so not $\vdash E$. Hence

$$\begin{aligned} (5) &\vdash \sim Thm^{\overline{"E"}} && \text{since } "E" \notin \mathcal{S} \\ (6) &\vdash E && .1, .5 \end{aligned}$$

This contradiction shows that \mathcal{A} cannot be decidable. ■

Note that if \mathcal{A} is inconsistent, then \mathcal{A} is decidable, since every wff_o is a theorem, and the set of numbers of wffs_o is recursive. Thus the consistency condition in the statement of the theorem above is clearly necessary. Note that this is a theorem about \mathcal{Q}_0^∞ , not \mathcal{Q}_0 . There are consistent decidable extensions of \mathcal{Q}_0 (see Exercise X8032). Also note that nothing was assumed about the axiomatizability of \mathcal{A} . Thus \mathcal{A} might have as axioms (and hence as theorems) the set of all sentences true in some particular model of \mathcal{Q}_0^∞ .

Since the theorem above is very deep and its proof is very short, it is worth taking time to reflect on the nature of the essential ingredients of the proof, so that one will know when one can apply the same basic idea in other contexts. First of all, it was essential that every decidable set could be represented in \mathcal{A} . To establish this, we used Church's Thesis and the fact that recursiveness implies representability in \mathcal{A} . Also, it was essential that we could construct a sentence \mathbf{E} such that $\vdash \mathbf{E} = \sim Thm_{o\sigma}^{\overline{"E"}}$. In the context of our Gödel numbering, this was accomplished by using the fact that * and Num are representable functions. Apart from these ingredients, the proof used only very elementary logic.

We can slightly generalize the result above. First we introduce some more terminology.

DEFINITION. A set \mathcal{H} of sentences is *consistent with* a logistic system \mathcal{S} iff not $\mathcal{H} \vdash_{\mathcal{S}} F_o$.

Note that \mathcal{H} is consistent with \mathcal{S} iff the result of adding the sentences in \mathcal{H} to the postulates of \mathcal{S} is a consistent system.

7301 Theorem. If \mathcal{B} is an extension of \mathcal{Q}_0 such that {Axiom 6} is consistent with \mathcal{B} , then \mathcal{B} is undecidable.

Proof: Let \mathcal{A} be the result of adding Axiom 6 to the postulates of \mathcal{B} . Thus \mathcal{A} is a consistent extension of \mathcal{Q}_0^∞ . If \mathcal{B} is not a *pure* extension of \mathcal{Q}_0 , \mathcal{A} will not be a pure extension of \mathcal{Q}_0^∞ , but the proof of 7300 can easily be modified to show that \mathcal{A} must be undecidable; one must simply introduce a suitable Gödel numbering for the wffs of \mathcal{B} . Note that the proof of 7300 depends primarily on the fact that every recursive function or relation is representable in \mathcal{Q}_0^∞ , and hence in every extension of \mathcal{Q}_0^∞ .

Now for each wff_o \mathbf{C} of \mathcal{B} , $\vdash_{\mathcal{A}} \mathbf{C}$ iff $\vdash_{\mathcal{B}} [\text{Axiom 6}] \supset \mathbf{C}$. Thus if \mathcal{B} were decidable, \mathcal{A} would be also. Hence \mathcal{B} must be undecidable. ■

7302 Corollary. \mathcal{Q}_0 is undecidable.

Proof: By 7301 and 7204. ■

DEFINITION. A logistic system \mathcal{A} is *complete* iff for each sentence \mathbf{B} , either $\vdash_{\mathcal{A}} \mathbf{B}$ or $\vdash_{\mathcal{A}} \sim \mathbf{B}$.

It is worth noting that the Gödel-Rosser Theorem (7200) follows easily from 7300 and the following proposition:

7303 Proposition. Let \mathcal{A} be an extension of \mathcal{Q}_0 (or of \mathcal{F}) which is recursively axiomatized and complete. Then \mathcal{A} is decidable.

Proof: If \mathcal{A} is inconsistent, it is clearly decidable, so we may assume that \mathcal{A} is consistent. It suffices to show that there is an effective procedure for deciding whether *sentences* of \mathcal{A} are theorems, since a wff is a theorem iff its universal closure is. Since for each sentence \mathbf{B} , either $\vdash_{\mathcal{A}} \mathbf{B}$ or $\vdash_{\mathcal{A}} \sim \mathbf{B}$, to determine whether \mathbf{B} is a theorem one can simply enumerate proofs (since \mathcal{A} is recursively axiomatized) until one finds a proof of \mathbf{B} or of $\sim \mathbf{B}$. ■

7200A Gödel-Rosser Incompleteness Theorem. Let \mathcal{A} be a consistent, recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Then \mathcal{A} is not complete.

Proof: By 7300 and 7303. ■

REMARK. While the proof above is very elegant and does not use theorems 6210 and 6212, it does not provide an explicit example of a sentence which is neither provable nor refutable.

We next turn our attention to the semantic notion of definability.

DEFINITIONS.

- (1) Let $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{J} \rangle$ be a general model for \mathcal{Q}_0 and let $r \subseteq \mathcal{D}_{\alpha_1} \times \dots \times \mathcal{D}_{\alpha_n}$, where $n \geq 1$. (Thus r is an n -ary relation on \mathcal{M} .) We say that r is a *definable* relation iff there is a cwff $\mathbf{B}_{o\alpha_n\dots\alpha_1}$ such that for all n -tuples $\langle k_1, \dots, k_n \rangle \in \mathcal{D}_{\alpha_1} \times \dots \times \mathcal{D}_{\alpha_n}$, $r(k_1, \dots, k_n)$ iff $\mathcal{M} \models \mathbf{B}_{o\alpha_n\dots\alpha_1} x_{\alpha_1}^1 \dots x_{\alpha_n}^n$, where $\varphi x_{\alpha_i}^i = k_i$ for $1 \leq i \leq n$. In this case we can clearly identify r with $\mathcal{V}^{\mathcal{M}} \mathbf{B}_{o\alpha_n\dots\alpha_1}$, so we can also say that r is definable iff there is a cwff $\mathbf{B}_{o\alpha_n\dots\alpha_1}$ such that $r = \mathcal{V}^{\mathcal{M}} \mathbf{B}_{o\alpha_n\dots\alpha_1}$.
- (2) If \mathcal{M} is any general model for \mathcal{Q}_0^∞ and r is any n -ary numerical relation, we say that r is *definable in \mathcal{M}* iff there is a cwff $\mathbf{B}_{o\sigma\dots\sigma}$ such that for all natural numbers k_1, \dots, k_n , $r(k_1, \dots, k_n)$ iff $\mathcal{M} \models \mathbf{B}_{o\sigma\dots\sigma} \bar{k}_1 \dots \bar{k}_n$. Note that if \mathcal{M} is any general model for \mathcal{Q}_0^∞ , we can identify the natural numbers with the finite cardinals in \mathcal{D}_σ , and $\mathcal{V}^{\mathcal{M}} \bar{k} = k$ for each natural number k , so this definition of a definable numerical relation can be regarded as a special case of the more general definition above.

7304 Proposition. If \mathcal{M} is any model for \mathcal{Q}_0^∞ , then every numerical relation which is representable in \mathcal{Q}_0^∞ is definable in \mathcal{M} .

Proof: Suppose r is represented in \mathcal{Q}_0^∞ by $\mathbf{B}_{o\sigma\dots\sigma}$. If $r(k_1, \dots, k_n)$, then $\vdash_{\mathcal{Q}_0^\infty} \mathbf{B} \bar{k}_1 \dots \bar{k}_n$, so $\mathcal{M} \models \mathbf{B} \bar{k}_1 \dots \bar{k}_n$. If not $r(k_1, \dots, k_n)$, then $\vdash_{\mathcal{Q}_0^\infty} \sim \mathbf{B} \bar{k}_1 \dots \bar{k}_n$, so $\mathcal{M} \models \sim \mathbf{B} \bar{k}_1 \dots \bar{k}_n$, so not $\mathcal{M} \models \mathbf{B} \bar{k}_1 \dots \bar{k}_n$. ■

However, certain definable relations are not representable. For example, we showed in 7300 that the set Thm of Gödel numbers of theorems of \mathcal{Q}_0^∞ is not recursive, and hence not representable. However, when we take \mathcal{A} to be \mathcal{Q}_0^∞ , the wff *Theorem* _{$o\sigma$} of §71 (namely $[\lambda x_\sigma \dot{\exists} y_\sigma \text{Proof}_{\sigma\sigma\sigma} x_\sigma y_\sigma]$) defines this set in any standard model \mathcal{M} of \mathcal{Q}_0^∞ . Indeed, if $\text{Thm}(k)$, then there is a wff _{σ} \mathbf{A} such that $k = "A"$ and $\vdash \mathbf{A}$, so $\vdash \text{Theorem}_{o\sigma} \bar{k}$ by 7103, so $\mathcal{M} \models \text{Theorem}_{o\sigma} \bar{k}$. On the other hand, if not $\text{Thm}(k)$, then for every number

$m, \vdash \sim Proof_{o\sigma\sigma}\overline{km}$, so $\mathcal{M} \models \sim Proof_{o\sigma\sigma}\overline{km}$, so $\mathcal{M} \models \dot{\forall}y_\sigma \sim Proof_{o\sigma\sigma}\overline{ky}_\sigma$ (since \mathcal{M} is a *standard* model), so $\mathcal{M} \models \sim Theorem_{o\sigma}\overline{k}$.

We next show that the set of Gödel numbers of sentences *true* in \mathcal{M} is not even definable in \mathcal{M} .

7305 Tarski's Theorem [Tarski, 1936]. For any model \mathcal{M} of \mathcal{Q}_0^∞ , $\{\text{"A"} \mid \mathbf{A} \text{ is a sentence and } \mathcal{M} \models \mathbf{A}\}$ is not definable in \mathcal{M} .

Proof: Suppose the given set \mathcal{S} is definable by the wff $\mathbf{B}_{o\sigma}$. Thus for each sentence \mathbf{A} , $\mathcal{M} \models \mathbf{A}$ iff $\text{"A"} \in \mathcal{S}$ iff $\mathcal{M} \models \mathbf{B}_{o\sigma}\overline{\text{"A"}}$. Let $\mathbf{C}_{o\sigma}$ be $[\lambda x_\sigma \bullet \sim \mathbf{B}_{o\sigma}[128 * x_\sigma * [\text{Num}_{o\sigma}x_\sigma] * 2048]]$, and let \mathbf{A}_o be $[\mathbf{C}_{o\sigma}\overline{\mathbf{C}_{o\sigma}}]$. Then $\vdash \mathbf{A} = \sim \mathbf{B}_{o\sigma}\overline{\text{"A"}}$ by the definition of \mathbf{A} , so $\mathcal{M} \models \mathbf{A}$ iff not $\mathcal{M} \models \mathbf{B}_{o\sigma}\overline{\text{"A"}}$. Since \mathbf{A} is a sentence, this is a contradiction. ■

REMARK. After proving 7303 we noted that 7200A (the Gödel-Rosser Theorem) follows directly from 7300 and 7303. Let us also see how to derive 7200A from Tarski's Theorem. Suppose \mathcal{A} is a consistent and complete recursively axiomatized pure extension of \mathcal{Q}_0^∞ . By 5501 \mathcal{A} has a model, which we shall call \mathcal{M} . Let $\mathcal{T} = \{\text{"A"} \mid \mathbf{A} \text{ is a sentence and } \mathcal{M} \models \mathbf{A}\}$. Since \mathcal{M} is a model of \mathcal{A} and \mathcal{A} is complete, it is clear that $\mathcal{M} \models \mathbf{A}$ iff $\vdash_{\mathcal{A}} \mathbf{A}$ for each sentence \mathbf{A} , so $\mathcal{T} = \{\text{"A"} \mid \mathbf{A} \text{ is a sentence and } \vdash_{\mathcal{A}} \mathbf{A}\}$. \mathcal{A} is decidable by 7303, so \mathcal{T} is recursive, and hence representable in \mathcal{Q}_0^∞ (by 6401), and hence definable in \mathcal{M} (by 7304). However, this contradicts Tarski's Theorem.

Let us also look at the problem of defining truth from a different perspective. When studying logic, we ordinarily try to distinguish carefully between syntactic and semantic concepts, and we only speak of a sentence as being true or false when we have a particular model in mind. However, one can also speak of truth in a more absolute sense, where one does not mention a model because one is concerned with only one model, the actual world. Let us now consider the problem of defining truth without discussing models.

Of course, questions about the nature of truth have arisen throughout the history of philosophy, so it is reasonable to ask how one is to know whether a proposed definition of truth is "adequate". Tarski has suggested [Tarski, 1936] that an adequate definition of truth for a language \mathcal{L}^1 which is given in a meta-language \mathcal{L}^2 of \mathcal{L}^1 should satisfy at least the following criterion: For each sentence \mathbf{A} of \mathcal{L}^1 , $\vdash_{\mathcal{L}^2} \mathbf{A} \equiv \text{True } \text{"A"}$, where "A" is the name in \mathcal{L}^2 for the sentence \mathbf{A} . For example, if one is using a formalized version of English as a meta-language for French (the object language), and

one has an adequate definition of truth for the object language, one should be able to prove that it is snowing iff the sentence “Il neige” is true. If the object language is itself English, one should be able to prove that it is snowing iff the sentence “It is snowing” is true. In [Andrews, 1965, pages 83 and 99] it is shown that a definition of truth for \mathcal{Q}_0 which satisfies Tarski’s criterion for an adequate definition of truth can be given in the system \mathcal{Q} of transfinite type theory.

To some extent a language may serve as its own meta-language. Indeed, we have shown that by using numerals representing Gödel numbers as names for formulas of \mathcal{Q}_0 , various syntactical notions for \mathcal{Q}_0^∞ such as ‘wff’ and ‘proof’ can be represented in \mathcal{Q}_0^∞ . It is therefore natural to ask whether the semantical notion ‘truth’ for \mathcal{Q}_0^∞ can be adequately defined within \mathcal{Q}_0^∞ . Tarski showed that the answer is negative.

7306 Theorem. Let \mathcal{A} be a consistent pure extension of \mathcal{Q}_0^∞ . There is no wff $\mathbf{B}_{o\sigma}$ of \mathcal{A} such that for each sentence \mathbf{A} of \mathcal{A} , $\vdash_{\mathcal{A}} \mathbf{A} = \mathbf{B}_{o\sigma} \overline{\text{“A”}}$.

Proof: Suppose there were such a wff $\mathbf{B}_{o\sigma}$. Then as in the proof of 7305 we can construct a sentence \mathbf{A} such that $\vdash_{\mathcal{A}} \mathbf{A} = \sim \mathbf{B}_{o\sigma} \overline{\text{“A”}}$. (If $\mathbf{B}_{o\sigma} \overline{\text{“C”}}$ says “‘C’ is true”, then \mathbf{A} says ‘I am not true’.) Thus $\vdash_{\mathcal{A}} \mathbf{B}_{o\sigma} \overline{\text{“A”}} = \sim \mathbf{B}_{o\sigma} \overline{\text{“A”}}$, so \mathcal{A} is inconsistent. ■

EXERCISES

In each exercise below, discuss whether there is an algorithm for determining of an arbitrary sentence \mathbf{A} whether or not it has the property in question. Justify your answers with proofs.

X7300. \mathbf{A} is true in all general models.

X7301. \mathbf{A} is true in all standard models.

X7302. \mathbf{A} is true in all infinite general models.

X7303. \mathbf{A} is true in all infinite standard models.

X7304. \mathbf{A} is satisfiable.

X7305. \mathbf{A} is satisfiable in some standard models.

§74. Epilogue

If one looks beyond the technical details of the theorems in this chapter and seeks to put them in philosophical perspective, one can summarize the main import of this chapter with the following conclusion:

Truth is Elusive.

The abstract concept of truth is at least partially manifested in the set of all statements (in whatever language) which are true; in particular, it is manifested in the set, which we call Truth , of all sentences of \mathcal{Q}_0 which are true in the actual existing physical and conceptual universe. By including concepts as well as physical objects as entities in our model of reality, we guarantee that it is infinite even if the physical universe and all aspects of time are finite. Thus, we may regard the actual universe as a model for \mathcal{Q}_0^∞ . (Of course, one might specify the domain of individuals in a variety of ways.) Since Truth is closed under the rules of logical inference, we may regard Truth not only as a set of sentences, but also as a consistent and complete extension of \mathcal{Q}_0^∞ whose axioms are the sentences in Truth , and whose theorems are the wffs_o whose universal closures are in Truth . Let us note how the results of this chapter imply that Truth is elusive:

- (1) By 7300, Truth is not decidable. That is, there is no correct and universally applicable algorithm for determining what is true, and what is not true.
- (2) By 7201a, Truth is not recursively axiomatizable. By 7200, if we have correct recursively axiomatized theories about what is true, they will inevitably be incomplete.
- (3) By 7305, Truth is not even definable (in the language in which the truths are expressed).

Supplementary Exercises

X8000. Find disjunctive and conjunctive normal forms for the wff
 $[(p \wedge q) \equiv \sim r \vee \sim q \wedge \sim p]$.

X8001. Express the following statements in symbolic form, and prove that the conclusion follows from the stated assertions.

- (a) Anyone who can fly impresses everyone.
- (b) Some people are not impressed by anyone.
- (c) Therefore, no one can fly.

X8002. For each of the wffs below, find an equivalent wff in which the scope of each negation is an atomic wff. Here we are using a formulation of first-order logic in which \sim , \wedge , \vee , \supset , \equiv , \forall , and \exists are all primitive symbols.

- (a) $\sim \sim [\forall x \exists y Mxy \vee \forall z \sim Qz] \supset \forall y Pyy \vee \exists z Rz$
- (b) $\sim \forall x [Px \equiv Qx] \equiv \exists x [Rx \equiv Qy]$
- (c) $\sim \sim [\exists x Qx \vee \forall x \sim \exists y Pxy] \supset \forall x Qx \wedge \exists y Pyy$
- (d) $\sim \sim [\sim \forall x \forall y Pxy \supset \forall z Qz] \equiv \exists x Qx$

X8003. Let A be the wff $\forall z Pxz \supset \exists x \exists y [Rx \vee Qy]$. Find a wff B such that the number of quantifiers occurring in B is as small as possible and $\vdash A \equiv B$. Give a sequence M_1, \dots, M_n of wffs such that M_1 is A , M_n is B , and one can readily see that $\vdash M_i \equiv M_{i+1}$ for each $i < n$ by applying theorems from the text.

X8004. For this exercise we assume we have a formulation of first-order logic in which \sim , \wedge , \vee , \supset , \forall , and \exists are all primitive symbols. Let D be the wff

$$\exists x \forall y \exists w \forall z \sim [Pxy \vee Qxy] \supset \sim Rzw \wedge Swz$$

Find a wff E which is equivalent to D such that each quantifier in E has a scope which contains as few occurrences of propositional connectives as possible. (The *number* of occurrences of quantifiers may be as large as you like.)

In exercises X8005–X8006, prove the given wffs. Present your proofs in the usual format, numbering each line, and indicating the hypotheses and justification for each line of a proof.

X8005. $\forall x \exists y [Px \wedge Qy] \supset \neg Qx \vee Py$

X8006. $\forall x \exists y [Px \wedge \neg Qy \vee Qx] \supset \exists z Pz \wedge Qz$

In exercises X8007–X8019, decide whether the given wff is valid or not.

X8007. $\exists x [Px \supset \forall y Py]$

X8008. $\exists x [Px \supset \exists y Py]$

X8009. $\forall x [Px \supset \forall y Py]$

X8010. $\forall x [Px \wedge Qx] \equiv \neg \forall x Px \wedge \forall x Qx$

X8011. $\forall x [Px \equiv Qx] \supset \neg \exists x Px \supset \exists x Qx$

X8012. $\exists x [Px \equiv Qx] \supset \neg \exists x Px \supset \exists x Qx$

X8013. $\forall y \exists z \exists w \forall x [\sim Px \vee Qy] \supset \neg Pz \supset Qw$

X8014. $\forall u \forall v \forall w [Ruv \vee Rvw] \supset \exists x \forall y Rxy$

X8015. $\forall x \exists y [Px \wedge Qy] \supset Qx$

X8016. $\forall x [Px \vee Qx] \equiv \neg \forall x Px \vee \forall x Qx$

X8017. $\forall x [Px \supset Qx] \supset \neg \exists x [Qx \supset Rx] \supset \exists x [Px \supset Rx]$

X8018. $\forall x [Px \supset Qx] \supset \neg \exists x [Px \supset Rx] \supset \exists x [Qx \supset Rx]$

X8019. $\exists w [\forall x Rxx \supset \exists z Rzw]$

X8020. Let D be the wff $\forall x \exists y Pxy \supset \exists w \forall z Pwz$.

- (a) Is D valid?
- (b) Let \mathcal{T} be the first-order theory obtained from the system \mathcal{F} of first-order logic by adding D as an additional axiom. Is \mathcal{T} consistent?

X8021. Let A be the wff $\forall x Rxy \supset \exists z \forall y [Rxy \wedge Ryz] \supset \sim \forall u Ru z$

- (a) Is x free for u in A ?
- (b) Is x free for y in A ?
- (c) Find a prenex normal form of A .

X8022. Give examples of theorems which show the relationship between syntactical concepts and semantic concepts.

X8023. Prove or refute the following conjecture: For every wff C of \mathcal{F} , $\models C$ if and only if $\models [C \supset \forall x C]$.

X8024. Let W be the sentence $\exists z_i [\lambda x_i z_i] = [\lambda x_i x_i]$.

- (a) Is W a theorem of Q_o ? Justify your answer.
- (b) Is $\{W\}$ a consistent set of sentences of Q_o ? Justify your answer.

X8025. In each case below, decide whether the given wff is a theorem of Q_o . Justify your answers with rigorous arguments.

- (a) $\forall f_n \forall g_n \exists z_i [f_n z_i = g_n z_i] \supset \forall x_i \forall y_i [x_i = y_i]$
- (b) $\exists p_{oi} \exists y_i \sim p_{oi} y_i \wedge \forall f_n \forall x_i p_{oi} f_n x_i$

X8026.

- (a) Explain, in an informal and intuitive way, why \sim Axiom 6 must be true in every finite general model.
- (b) Show that there is a general model \mathcal{M} for Q_0 with an infinite domain of individuals such that $\mathcal{M} \models \sim$ Axiom 6.

X8027. Give an example of a set of natural numbers which is not representable in Q_o^∞ .

X8028. Let \mathcal{H} be the set of all sentences which are true in all standard models. Does \mathcal{H} have any nonstandard models?

X8029. In this exercise, the object language is Q_o^∞ . Let $Proof_{o\sigma\sigma}$ be a wff which represents the numerical relation Proof, and $Theorem_{o\sigma}$ be $[\lambda x_\sigma \dot{\exists} y_\sigma \cdot Proof_{o\sigma\sigma} x_\sigma y_\sigma]$. Let $M_{o\sigma\sigma}$ be the wff

$$[\lambda x_\sigma \lambda y_\sigma \cdot Theorem_{o\sigma} [\overline{["}] * y_\sigma * [Num_{\sigma\sigma} x_\sigma] * \overline{"]}] * [Num_{\sigma\sigma} y_\sigma] * \overline{"]}]$$

and let $N_{o\sigma\sigma}$ be the wff

$$[\lambda x_\sigma \lambda y_\sigma \cdot Theorem_{o\sigma} [\overline{["}] * x_\sigma * [Num_{\sigma\sigma} x_\sigma] * \overline{"]}] * [Num_{\sigma\sigma} y_\sigma] * \overline{"]}].$$

Let $m = "M_{o\sigma\sigma}"$ and $n = "N_{o\sigma\sigma}"$.

Let $D = [[M_{o\sigma\sigma} \overline{m}] \overline{n}]$ and let $E = [[N_{o\sigma\sigma} \overline{m}] \overline{n}]$.

- (a) Prove that $\vdash D$ if and only if $\vdash E$.
- (b) Is D a theorem? Justify your answer.

X8030. We start by defining the system \mathcal{C} , which has the same type symbols, improper symbols, variables, and formation rules as Q_0 , but has the following logical constants: $\sim_{(oo)} \vee_{((oo)o)} \Pi_{(o(o\alpha))} \iota_{(\iota(o\alpha))}$

Definitions and abbreviations for C:

1. $[\mathbf{A}_o \vee \mathbf{B}_o]$ stands for $[[\vee_{((oo)o)} \mathbf{A}_o] \mathbf{B}_o]$.
2. $[\mathbf{A}_o \supset \mathbf{B}_o]$ stands for $[[\sim_{oo} \mathbf{A}_o] \vee \mathbf{B}_o]$.
3. $[\forall \mathbf{x}_\alpha \mathbf{A}_o]$ stands for $[\Pi_{(o(o\alpha))} [\lambda \mathbf{x}_\alpha \mathbf{A}_o]]$.
4. Other propositional connectives, and the existential quantifier, are defined in familiar ways.
5. $\mathcal{Q}_{o\alpha}$ stands for $[\lambda x_\alpha \lambda y_\alpha \forall f_{o\alpha} \cdot f_{o\alpha} x_\alpha \supset f_{o\alpha} y_\alpha]$.
6. $[\mathbf{A}_\alpha = \mathbf{B}_\alpha]$ stands for $\mathcal{Q}_{o\alpha} \mathbf{A}_\alpha \mathbf{B}_\alpha$.
7. $\exists_1 \mathbf{x}_\alpha \mathbf{A}_o$ stands for
 $[\lambda p_{o\alpha} \cdot \exists y_\alpha \cdot p_{o\alpha} y_\alpha \wedge \forall z_\alpha \cdot p_{o\alpha} z_\alpha \supset z_\alpha = y_\alpha] [\lambda \mathbf{x}_\alpha \mathbf{A}_o]$.

Rules of inference of C:

1. *Alphabetic Change of Bound Variables* (α -conversion). To replace any well-formed part $[\lambda \mathbf{x}_\beta \mathbf{A}_\alpha]$ of a wff by $[\lambda \mathbf{y}_\beta S_{\mathbf{y}_\beta}^{\mathbf{x}_\beta} \mathbf{A}_\alpha]$, provided that \mathbf{y}_β does not occur in \mathbf{A}_α and \mathbf{x}_β is not bound in \mathbf{A}_α .
2. *β -contraction.* To replace any well-formed part $[[\lambda \mathbf{x}_\alpha \mathbf{B}_\beta] \mathbf{A}_\alpha]$ of a wff by $S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} \mathbf{B}_\beta$, provided that the bound variables of \mathbf{B}_β are distinct both from \mathbf{x}_α and from the free variables of \mathbf{A}_α .
3. *β -expansion.* To infer \mathbf{C} from \mathbf{D} if \mathbf{D} can be inferred from \mathbf{C} by a single application of β -contraction.
4. *Substitution.* From $\mathbf{F}_{o\alpha} \mathbf{x}_\alpha$, to infer $\mathbf{F}_{o\alpha} \mathbf{A}_\alpha$, provided that \mathbf{x}_α is not a free variable of $\mathbf{F}_{o\alpha}$.
5. *Modus Ponens.* From $[\mathbf{A}_o \supset \mathbf{B}_o]$ and \mathbf{A}_o , to infer \mathbf{B}_o .
6. *Generalization.* From $\mathbf{F}_{o\alpha} \mathbf{x}_\alpha$ to infer $\Pi_{o(o\alpha)} \mathbf{F}_{o\alpha}$, provided that \mathbf{x}_α is not a free variable of $\mathbf{F}_{o\alpha}$.

Axioms of C:

- (1) $p_o \vee p_o \supset p_o$
- (2) $p_o \supset p_o \vee q_o$

$$(3) p_o \vee q_o \supset q_o \vee p_o$$

$$(4) p_o \supset q_o \supset [r_o \vee p_o \supset r_o \vee q_o]$$

$$(5^\alpha) \Pi_{o(o\alpha)} f_{o\alpha} \supset f_{o\alpha} x_\alpha$$

$$(6^\alpha) \forall x_\alpha [p_o \vee f_{o\alpha} x_\alpha] \supset \neg p_o \vee \Pi_{o(o\alpha)} f_{o\alpha}$$

(7) Axioms of Extensionality:

$$(7^o) [x_o \equiv y_o] \supset x_o = y_o$$

$$(7^{\alpha\beta}) \forall x_\beta [f_{\alpha\beta} x_\beta = g_{\alpha\beta} x_\beta] \supset f_{\alpha\beta} = g_{\alpha\beta}$$

(8) Axiom of Descriptions:

$$\exists_1 y_\alpha p_{o\alpha} y_\alpha \supset p_{o\alpha} [\iota_{o(o\alpha)} p_{o\alpha}]$$

\mathcal{C} is essentially the system introduced by Church in [Church, 1940], except that Church's system did not have Axiom 7^o and did have axioms of infinity and choice. The system proved complete by Henkin in [Henkin, 1950] is obtained when Axiom 8 of \mathcal{C} is replaced by an axiom of choice.

A notable subsystem \mathcal{T} of \mathcal{C} is obtained by deleting Axioms (7^o) , $(7^{\alpha\beta})$, and (8). \mathcal{T} may also be called *elementary type theory* [Andrews, 1971][Andrews, 1974a], since \mathcal{T} simply embodies the logic of propositional connectives, quantifiers, and λ -conversion in the context of type theory.

Show that \mathcal{C} and \mathcal{Q}_0 are equivalent systems in the sense that there are type-preserving functions σ mapping wffs of \mathcal{C} to wffs of \mathcal{Q}_0 and τ mapping wffs of \mathcal{Q}_0 to wffs of \mathcal{C} such that

- (a) For every wff \mathbf{A}_o of \mathcal{C} , if $\vdash_{\mathcal{C}} \mathbf{A}_o$, then $\vdash_{\mathcal{Q}_0} \sigma \mathbf{A}_o$.
- (b) For every wff \mathbf{B}_o of \mathcal{Q}_0 , if $\vdash_{\mathcal{Q}_0} \mathbf{B}_o$, then $\vdash_{\mathcal{C}} \tau \mathbf{B}_o$.
- (c) For every wff \mathbf{A} of \mathcal{C} , $\vdash_{\mathcal{C}} \mathbf{A} = \tau(\sigma \mathbf{A})$.
- (d) For every wff \mathbf{B} of \mathcal{Q}_0 , $\vdash_{\mathcal{Q}_0} \mathbf{B} = \sigma(\tau \mathbf{B})$.

X8031. Prove the following theorem of \mathcal{Q}_o :

$$\text{Finite}_{o(o)} [\lambda x_i \neg x_i = y_i \vee x_i = z_i]$$

X8032. Let \mathcal{Q}_0^1 be the system obtained by adding to \mathcal{Q}_0 the additional axiom $\forall x_i \forall y_i \neg x_i = y_i$. Prove that \mathcal{Q}_0^1 is consistent, complete, and decidable. (*Hint:* See Exercise X5502.)

Summary of Theorems

1000 Principle of Induction on the Construction of a Wff. Let \mathcal{R} be a property of formulas, and let $\mathcal{R}(\mathbf{A})$ mean that \mathbf{A} has property \mathcal{R} . Suppose

- (1) $\mathcal{R}(\mathbf{p})$ for each propositional variable \mathbf{p} .
- (2) Whenever $\mathcal{R}(\mathbf{A})$, then $\mathcal{R}(\sim \mathbf{A})$.
- (3) Whenever $\mathcal{R}(\mathbf{A})$ and $\mathcal{R}(\mathbf{B})$, then $\mathcal{R}([\mathbf{A} \vee \mathbf{B}])$.

Then every wff has property \mathcal{R} .

1100. If $\mathcal{H}_1 \vdash \mathbf{A}$ and $\mathcal{H}_1 \subseteq \mathcal{H}_2$, then $\mathcal{H}_2 \vdash \mathbf{A}$.

Corollary. If $\mathcal{H}_1 \vdash \mathbf{A}$ and $\mathcal{H}_2 \vdash \mathbf{A} \supset \mathbf{B}$ and $\mathcal{H}_1 \subseteq \mathcal{H}$ and $\mathcal{H}_2 \subseteq \mathcal{H}$, then $\mathcal{H} \vdash \mathbf{B}$.

1101 Rule of Substitution (Sub). If $\mathcal{H} \vdash \mathbf{A}$, and if $\mathbf{p}_1, \dots, \mathbf{p}_n$ are distinct variables which do not occur in any wff in \mathcal{H} , then $\mathcal{H} \vdash S_{\mathbf{B}_1 \dots \mathbf{B}_n}^{\mathbf{p}_1 \dots \mathbf{p}_n} \mathbf{A}$.

1102. If $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}$ and $\mathcal{H} \vdash \mathbf{C} \vee \mathbf{A}$, then $\mathcal{H} \vdash \mathbf{B} \vee \mathbf{C}$.

1103. $\vdash p \vee \sim p$

1104. $\vdash p \supset \sim \sim p$

1105. $\vdash \sim \sim p \supset p$

1106. $\vdash p \supset p$

1107. $\vdash q \vee p \supset p \vee q$

1108. If $\mathcal{H} \vdash \mathbf{A} \vee \mathbf{B}$, then $\mathbf{B} \vee \mathbf{A}$.

1109 Transitive Law of Implication (Trans). If $\mathcal{H} \vdash \mathbf{A}_1 \supset \mathbf{A}_2$, $\mathcal{H} \vdash \mathbf{A}_2 \supset \mathbf{A}_3, \dots$, and $\mathcal{H} \vdash \mathbf{A}_{n-1} \supset \mathbf{A}_n$, then $\mathcal{H} \vdash \mathbf{A}_1 \supset \mathbf{A}_n$.

1110. If $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{C}$ and $\mathcal{H} \vdash \mathbf{B} \supset \mathbf{C}$, then $\mathcal{H} \vdash \mathbf{A} \vee \mathbf{B} \supset \mathbf{C}$.

1111. If $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{C}$ and $\mathcal{H} \vdash \sim \mathbf{A} \supset \mathbf{C}$, then $\mathcal{H} \vdash \mathbf{C}$.

1112. If $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}$, then $\mathcal{H} \vdash \mathbf{A} \supset \neg \mathbf{C} \vee \mathbf{B}$ and $\mathcal{H} \vdash \mathbf{A} \supset \neg \mathbf{B} \vee \mathbf{C}$.

1113. $\vdash [p \vee q] \vee r \supset \neg p \vee \neg q \vee r$

1114. If $\mathcal{H} \vdash [\mathbf{A} \vee \mathbf{B}] \vee \mathbf{C}$, then $\mathcal{H} \vdash \mathbf{A} \vee \neg \mathbf{B} \vee \mathbf{C}$.

1115. If $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}$ and $\mathcal{H} \vdash \mathbf{A} \supset \neg \mathbf{B} \supset \mathbf{C}$, then $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{C}$.

1116 Deduction Theorem. If $\mathcal{H}, \mathbf{A} \vdash \mathbf{B}$, then $\mathcal{H} \vdash \mathbf{A} \supset \mathbf{B}$.

1117. $\vdash \neg p \supset \neg \neg q \supset \neg \neg p \vee q$

1118. $\vdash p \supset \neg \neg p \supset q$

1119. Principle of Induction on Proofs. Let \mathcal{L} be any logistic system, and let \mathcal{R} be any property of wffs of \mathcal{L} . If every axiom of \mathcal{L} has property \mathcal{R} , and each rule of inference of \mathcal{L} preserves property \mathcal{R} , then every theorem of \mathcal{L} has property \mathcal{R} .

1200 Soundness Theorem. Every theorem of \mathcal{P} is a tautology.

1201 Theorem. Let \mathcal{S} be any logistic system in which $\mathbf{A} \supset \neg \mathbf{A} \supset \mathbf{B}$ is a theorem schema and in which Modus Ponens is a primitive or derived rule of inference. Then \mathcal{S} is absolutely consistent iff \mathcal{A} is consistent with respect to negation.

1202 Consistency Theorem. \mathcal{P} is absolutely consistent and consistent with respect to negation.

1203 Lemma. Let \mathbf{A} be a wff and let all the variables in \mathbf{A} be among p_1, \dots, p_n . Let φ be any assignment. For any wff \mathbf{B} , let

$$\begin{aligned}\mathbf{B}^\varphi &= \mathbf{B} \text{ if } \mathcal{V}_\varphi \mathbf{B} = \top \\ &= \neg \mathbf{B} \text{ if } \mathcal{V}_\varphi \mathbf{B} = \mathbb{F}.\end{aligned}$$

Then $\mathbf{p}_1^\varphi, \dots, \mathbf{p}_n^\varphi \vdash \mathbf{A}^\varphi$.

1204 Completeness Theorem. Every tautology is a theorem of \mathcal{P} .

1205 Theorem. A wff is a theorem of \mathcal{P} if and only if it is a tautology.

1300 Theorem. The axiom schemata and the rule of inference of \mathcal{P} are all independent.

1400 Substitution-Value Theorem. Let φ be any assignment and θ be any proper substitution. If \mathbf{A} is any wff of propositional calculus, then $\mathcal{V}_\varphi \theta \mathbf{A} = \mathcal{V}_{\varphi \circ \theta} \mathbf{A}$.

1401 Corollary If $\models \mathbf{A}$, and θ is any proper substitution, then $\models \theta \mathbf{A}$.

1402 Substitutivity of Equivalence. If \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} are wffs of propositional calculus, and \mathbf{D} is obtained from \mathbf{C} by replacing zero or more occurrences of \mathbf{A} in \mathbf{C} by occurrences of \mathbf{B} , and $\models [\mathbf{A} \equiv \mathbf{B}]$, then $\models [\mathbf{C} \equiv \mathbf{D}]$.

1403 Theorem. Let h be any truth function with n arguments, where $n \geq 1$, and let $\mathbf{p}_1, \dots, \mathbf{p}_n$ be distinct propositional variables. Then there is a wff \mathbf{A} in disjunctive normal form such that $h = [\lambda \mathbf{p}_1, \dots, \lambda \mathbf{p}_n \mathbf{A}]$.

1404 Corollary. Every wff of propositional calculus has a disjunctive normal form.

1405 Theorem. The following sets of propositional connectives are complete:

- | | |
|-------------------------------|------------------------------------|
| (a) $\{\vee, \sim\}$ | (e) $\{ \}$ |
| (b) $\{\wedge, \sim\}$ | (f) $\{\downarrow\}$ |
| (c) $\{\supset, \sim\}$ | (g) $\{\wedge, \not\equiv, \top\}$ |
| (d) $\{\supset, \mathbf{f}\}$ | |

1406 Theorem. Let \mathbf{A} be a wff in negation normal form, and let φ be an assignment. Then

- (a) φ satisfies \mathbf{A} iff there is a path $P \in \mathcal{VP}(\mathbf{A})$ such that φ satisfies every literal in P .
- (b) φ falsifies \mathbf{A} iff there is a path $P \in \mathcal{HP}(\mathbf{A})$ such that φ falsifies every literal in P .

1407 Theorem. Let \mathbf{A} be a wff in negation normal form. Then

- (a) $\bigvee_{P \in \mathcal{VP}(\mathbf{A})} \bigwedge P$ is a disjunctive normal form of \mathbf{A} .
- (b) $\bigwedge_{P \in \mathcal{HP}(\mathbf{A})} \bigvee P$ is a conjunctive normal form of \mathbf{A} .

1408 Theorem. Let \mathbf{A} be a wff in negation normal form. Then

- (a) \mathbf{A} is a contradiction iff every path in $\mathcal{VP}(\mathbf{A})$ contains complementary literals.

- (b) \mathbf{A} is a tautology iff every path in $\mathcal{HP}(\mathbf{A})$ contains complementary literals.

1500 Proposition. Let \mathcal{H} be a finite set of wffs of \mathcal{P} . \mathcal{H} is satisfiable iff \mathcal{H} is consistent in \mathcal{P} .

1501 Compactness Theorem. Let \mathcal{S} be an infinite set of wffs of propositional calculus such that every finite subset of \mathcal{S} is satisfiable. Then \mathcal{S} is satisfiable.

1502 Corollary. Let \mathcal{S} be any set of wffs of \mathcal{P} . \mathcal{S} is satisfiable iff \mathcal{S} is consistent in \mathcal{P} .

1600 Ground Resolution Theorem. Let \mathcal{C} be any set of clauses. Then \mathcal{C} is contradictory iff $\mathcal{C} \vdash \square$.

2100 Rule P. If \mathbf{B} is tautologous, then $\vdash \mathbf{B}$. If $\vdash \mathbf{A}_1, \dots, \vdash \mathbf{A}_n$, and if $[\mathbf{A}_1 \wedge \dots \wedge \mathbf{A}_n \supset \mathbf{B}]$ is tautologous, then $\vdash \mathbf{B}$.

2101 Consistency Theorem. \mathcal{F} is absolutely consistent and consistent with respect to negation.

2102. $\vdash \forall x_n \dots \forall x_1 \mathbf{A} \supset \mathbf{A}$.

2103. $\supset \forall$ Rule: If $\vdash \mathbf{A} \supset \mathbf{B}$ and x is not free in \mathbf{A} , then $\vdash \mathbf{A} \supset \forall x \mathbf{B}$.

2104. $\vdash \forall x [\mathbf{A} \supset \mathbf{B}] \supset \neg \forall x \mathbf{A}$, then $\vdash \mathbf{A} \supset \forall x \mathbf{B}$.

2105 Substitutivity of Implication. Let \mathbf{M} , \mathbf{N} , and \mathbf{A} be wffs, and let \mathbf{A}' be the result of replacing \mathbf{M} by \mathbf{N} at zero or more occurrences (henceforth called designated occurrences) of \mathbf{M} in \mathbf{A} . Let y_1, \dots, y_k be a list including all individual variables which occur free in \mathbf{M} or \mathbf{N} , but are bound in \mathbf{A} by quantifiers whose scopes contain designated occurrences of \mathbf{M} .

If the designated occurrences of \mathbf{M} are all positive in \mathbf{A} , then

$$(a) \vdash \forall y_1 \dots \forall y_k [\mathbf{M} \supset \mathbf{N}] \supset \neg \mathbf{A} \supset \mathbf{A}', \quad \text{and}$$

$$(b) \text{ if } \vdash \mathbf{M} \supset \mathbf{N}, \text{ then } \vdash \mathbf{A} \supset \mathbf{A}'.$$

If the designated occurrences of \mathbf{M} are all negative in \mathbf{A} , then

$$(c) \vdash \forall y_1 \dots \forall y_k [\mathbf{M} \supset \mathbf{N}] \supset \neg \mathbf{A}' \supset \mathbf{A}, \quad \text{and}$$

(d) if $\vdash M \supset N$, then $\vdash A' \supset A$.

2106 Substitutivity of Equivalence. Let M, N, A, A' , and y_1, \dots, y_k be as in 2105.

(a) $\vdash \forall y_1 \dots \forall y_k [M \equiv N] \supset A \equiv A'$

(b) If $\vdash M \equiv N$ then $\vdash A \equiv A'$.

(c) If $\vdash M \equiv N$ and $\vdash A$ then $\vdash A'$.

2107. $\vdash \forall x [M \equiv N] \supset \forall x M \equiv \forall x N$

2108. $\vdash \forall x [M \equiv N] \supset \exists x M \equiv \exists x N$

2109. $\vdash \forall x C \equiv \forall y S_y^x C$, provided that y is not free in C and y is free for x in C .

2110 Rule of Alphabetic Change of Bound Variables (α Rule). Suppose that A' is obtained from A upon replacing an occurrence of $\forall x C$ in A by an occurrence of $\forall y S_y^x C$, where y is not free in C and y is free for x in C . If $\vdash A$, then $\vdash A'$.

2111 Universal Instantiation (VI) If $\mathcal{H} \vdash \forall x A$, then $\mathcal{H} \vdash S_t^x A$, provided that t is a term free for x in A .

2112 Lemma. Suppose $\mathcal{H} \vdash A$, and let \mathcal{U} be a finite set of individual variables which are not free in A or in \mathcal{H} . Then there is a proof of A from \mathcal{H} in which there is no application of the rule of Generalization involving generalization on a member of \mathcal{U} .

2113 Theorem. If $\mathcal{H}_1 \vdash A$ and $\mathcal{H}_1 \subseteq \mathcal{H}_2$, then $\mathcal{H}_2 \vdash A$.

Corollary. If $\vdash A$, then $\mathcal{H} \vdash A$.

2114 Extended Rule P. If $\mathcal{H} \vdash A_1, \dots, \mathcal{H} \vdash A_n$, and if $[A_1 \wedge \dots \wedge A_n \supset B]$ is tautologous, then $\mathcal{H} \vdash B$. Also, if B is tautologous, $\mathcal{H} \vdash B$.

2115 Extended $\supset \forall$ Rule. If $\mathcal{H} \vdash A \supset B$ and if x is not free in \mathcal{H} or in A , then $\mathcal{H} \vdash A \supset \forall x B$.

2116 Deduction Theorem (Ded). If $\mathcal{H}, A \vdash B$ then $\mathcal{H} \vdash A \supset B$.

2117 Extended Substitutivity of Implication and Equivalence. Let M , N , and A be wffs, and let A' be the result of replacing M by N at certain designated occurrences of M in A . If $\mathcal{H} \vdash A$, then $\mathcal{H} \vdash A'$, provided that one of the following conditions is satisfied:

- (a) $\vdash M \supset N$, and the designated occurrences of M are all positive in A ;
- (b) $\vdash N \supset M$, and the designated occurrences of M are all negative in A ;
- (c) $\vdash M \equiv N$.

2118 Rule of Substitution (Sub). Let m be an individual variable and t be a term, or let m be a propositional variable and let t be a wff. If $\mathcal{H} \vdash A$, then $\mathcal{H} \vdash S_t^m A$, provided that m does not occur free in \mathcal{H} and t is free for m in A .

$$2119. \vdash \sim \exists x A \equiv \forall x \sim A$$

$$2120. \vdash \sim \forall x A \equiv \exists x \sim A$$

$$2121. \vdash \forall x \forall y A \equiv \forall y \forall x A$$

$$2122. \vdash \exists x \exists y A \equiv \exists y \exists x A$$

$$2123. \vdash (S_t^x A) \supset \exists x A \text{ provided that } t \text{ is a term free for } x \text{ in } A.$$

$$2124. \vdash A \supset \exists x A$$

$$2125. \vdash \forall x A \supset \exists x A$$

2126 Rule of Existential Generalization (\exists Gen). If $\mathcal{H} \vdash S_t^x A$, where t is a term free for x in A , then $\mathcal{H} \vdash \exists x A$.

$$2127. \vdash \forall x A \equiv A \text{ if } x \text{ is not free in } A.$$

$$2128. \vdash \exists x A \equiv A \text{ if } x \text{ is not free in } A.$$

$$2129. \vdash \forall x [A \wedge B] \equiv \neg \forall x A \wedge \forall x B$$

$$2130. \vdash \exists x [A \vee B] \equiv \neg \exists x A \vee \exists x B$$

$$2131. \vdash \exists x [A \vee B] \equiv \neg A \vee \exists x B \text{ if } x \text{ is not free in } A. \\ \vdash \exists x [B \vee A] \equiv \neg \exists x B \vee A \text{ if } x \text{ is not free in } A.$$

2132. $\vdash \forall x A \vee \forall x B \supset \forall x . A \vee B.$

2133. $\vdash \forall x [A \vee B] \equiv . A \vee \forall x B$ if x is not free in A .
 $\vdash \forall x [B \vee A] \equiv . \forall x B \vee A$ if x is not free in A .

2134. $\vdash \forall x [B \supset A] \equiv . \exists x B \supset A$ if x is not free in A .

2135 Existential Rule (\exists Rule). If $\mathcal{H}, B \vdash A$ and if x is not free in A or in \mathcal{H} , then $\mathcal{H}, \exists x B \vdash A$.

2136 Rule C. If $\mathcal{H} \vdash \exists x B$ and $\mathcal{H}, S_y^x B \vdash A$, where y is an individual variable which is free for x in B and which is not free in \mathcal{H} , $\exists x B$, or in A , then $\mathcal{H} \vdash A$.

2137. $\vdash \exists x \forall z A \supset \forall z \exists x A.$

2138. $\vdash \exists x [A \wedge B] \supset . \exists x A \wedge \exists x B$

2139 Rule of Cases. If $\mathcal{H} \vdash A \vee B$ and $\mathcal{H}, A \vdash C$ and $\mathcal{H}, B \vdash C$, then $\mathcal{H} \vdash C$.

2140 Corollary. If $\mathcal{H}, A \vdash C$, and $\mathcal{H}, \sim A \vdash C$ then $\mathcal{H} \vdash C$.

2141 Indirect Proof (IP). If $\mathcal{H}, \sim A \vdash B$, and $\mathcal{H}, \sim A \vdash \sim B$, then $\mathcal{H} \vdash A$. Also, if $\mathcal{H}, \sim A \vdash B \wedge \sim B$, then $\mathcal{H} \vdash A$.

2200 Prenex Normal Form Theorem. Every wff has a prenex normal form. If D is a rectified wff and G is the prenex normal form of D , then $\vdash D \equiv G$.

2300 Proposition. Let \mathcal{J} be an interpretation, Z a term or wff, and φ and ψ assignments into \mathcal{J} which agree on all variable which occur free in Z . Then $\mathcal{V}_\varphi^{\mathcal{J}} Z = \mathcal{V}_\psi^{\mathcal{J}} Z$.

2301 Proposition. Let A be a wff and M be an interpretation.

- (a) A is valid [in M] iff $\sim A$ is not satisfiable [in M].
- (b) A is satisfiable [in M] iff $\sim A$ is not valid [in M].
- (c) A is valid [in M] iff $\forall x A$ is valid [in M].
- (d) A is satisfiable [in M] iff $\exists x A$ is satisfiable [in M].

2302 Substitution-Value Theorem. If A is a variable, constant, term, or wff, and θ is a proper substitution for free occurrences of variables such that θ is free on A , and φ is any assignment, then $\mathcal{V}_\varphi \theta A = \mathcal{V}_{\varphi \circ \theta} A$.

2303 Soundness Theorem.

- (a) Every theorem of \mathcal{F} is valid.
- (b) If \mathcal{T} is a first-order theory and \mathcal{M} is a model of \mathcal{T} , then every theorem of \mathcal{T} is valid in \mathcal{M} .
- (c) If \mathcal{G} is a set of wffs and φ simultaneously satisfies \mathcal{G} (in some interpretation) and $\mathcal{G} \vdash \mathbf{B}$, then φ satisfies \mathbf{B} .

2304 Theorem. Every set of sentences which has a model is consistent.

2401 Theorem. The rules of inference and axiom schemata of \mathcal{F} are independent.

2500 Proposition. Every wff of \mathcal{F} is a literal or is of one of the forms $[\mathbf{A} \vee \mathbf{B}]$, $\forall x\mathbf{A}$, $\sim\sim\mathbf{A}$, $\sim[\mathbf{A} \vee \mathbf{B}]$, $\sim\forall x\mathbf{A}$.

2501 Lemma. Let \mathcal{G} be any set of wffs. $\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$ is inconsistent with \mathcal{G} iff $\mathcal{G} \vdash \sim\mathbf{A}_1 \vee \dots \vee \sim\mathbf{A}_n$. The wff \mathbf{A} is inconsistent with \mathcal{G} iff $\mathcal{G} \vdash \sim\mathbf{A}$.

2502 Proposition. Let \mathcal{F}^1 and \mathcal{F}^2 be first order theories, with \mathcal{F}^2 a conservative extension of \mathcal{F}^1 . Then \mathcal{F}^2 is consistent iff \mathcal{F}^1 is consistent.

2503 Proposition. Let \mathcal{F}^1 and \mathcal{F}^2 be formulations of \mathcal{F} , so that \mathcal{F}^2 is an expansion of \mathcal{F}^1 obtained by adding additional individual constants to the set of constants of \mathcal{F}^1 . Let \mathcal{G} be any set of sentences of \mathcal{F}^1 . Then

- (1) $\mathcal{F}^2 \cup \mathcal{G}$ is a conservative extension of $\mathcal{F}^1 \cup \mathcal{G}$.
- (2) \mathcal{G} is consistent in \mathcal{F}^1 iff it is consistent in \mathcal{F}^2 .

2504 Lemma. If Γ is a class of sets which is of finite character, then Γ is closed under subsets.

2505 Proposition. $\{ \mathcal{S} \mid \mathcal{S} \text{ is a consistent set of sentences of } \mathcal{F} \}$ is an abstract consistency class.

2506 Proposition. Let Γ be an abstract consistency class. Let $\Delta = \{\mathcal{S} \mid \text{every finite subset of } \mathcal{S} \text{ is in } \Gamma\}$. Then $\Gamma \subseteq \Delta$ and Δ is an abstract consistency class of finite character.

2507 Smullyan's Unifying Principle. If Γ is an abstract consistency class and \mathcal{S} is a sufficiently pure set of sentences and $\mathcal{S} \in \Gamma$, then \mathcal{S} has a frugal model.

2508 Theorem. Every consistent set of sentences has a frugal model.

2509 Generalized Completeness Theorem. If \mathcal{S} is a set of sentences and \mathbf{A} is valid in all frugal models of \mathcal{S} , then $\mathcal{S} \vdash \mathbf{A}$.

2510 Gödel's Completeness Theorem. Every valid wff of \mathcal{F} is a theorem.

2511 Löwenheim's Theorem. If \mathbf{A} is valid in every countable interpretation, then \mathbf{A} is valid.

2512 Theorem. A set of sentences is consistent iff it has a model.

2513 Löwenheim-Skolem Theorem. Every set of sentences which has a model has a frugal model.

2514 Compactness Theorem. A set \mathcal{S} of sentences has a model iff every finite subset of \mathcal{S} has a model.

2515 Proposition. Let \mathcal{F}^1 and \mathcal{F}^2 be formulations of \mathcal{F} , so that \mathcal{F}^2 is an expansion of \mathcal{F}^1 obtained by adding additional constants (of any type) to the set of constants of \mathcal{F}^1 . Let \mathcal{G} be any set of sentences of \mathcal{F}^1 . Then

- (1) $\mathcal{F}^2 \cup \mathcal{G}$ is a conservative extension of $\mathcal{F}^1 \cup \mathcal{G}$.
- (2) \mathcal{G} is consistent in \mathcal{F}^1 iff it is consistent in \mathcal{F}^2 .

2600. $\vdash x = y \supset y = x$.

2601. $\vdash x = y \supset \neg y = z \supset x = z$

2602. $\vdash x = y \supset \neg (S_x^z A) \equiv S_y^z A$, where x and y are free for z in A .

2603. $\vdash x = y \supset (S_x^z t) = (S_y^z t)$ for any term t .

2604. $\vdash x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset f x_1 \dots x_n = f y_1 \dots y_n$ for any n -ary function symbol f .

2605. $\vdash x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \neg P x_1 \dots x_n \equiv P y_1 \dots y_n$ for any n -ary predicate symbol P .

2606 Proposition. Let \mathcal{T} be a first-order theory with a binary predicate constant $=$ such that

$$(a) \vdash_{\mathcal{T}} x = x$$

and for each n -ary function symbol \mathbf{f}^n

$$(b) \vdash_{\mathcal{T}} x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \mathbf{f}^n x_1 \dots x_n = \mathbf{f}^n y_1 \dots y_n$$

and for each n -ary predicate symbol \mathbf{P}^n

$$(c) \vdash_{\mathcal{T}} x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \bullet \mathbf{P}_1^n \dots x_n \supset \mathbf{P}^n y_1 \dots y_n.$$

Then \mathcal{T} is a first-order theory with equality.

$$2607. \vdash \exists_1 x Px \equiv \exists x \bullet Px \wedge \forall y \bullet Py \supset x = y$$

$$2608. \vdash \exists_1 x Px \equiv \exists y \forall x \bullet Px \equiv x = y$$

2609 Soundness Theorem for Logic with Equality. If \mathcal{M} is an equality-model for a set \mathcal{S} of sentences and $\mathcal{S} \vdash_{\mathcal{F}^=} \mathbf{A}$, then $\mathcal{M} \models \mathbf{A}$.

2610 Proposition. Let \mathcal{M} be any model for Axioms 6 and 7. Then there is an equality-model \mathcal{N} for Axioms 6 and 7 such that the cardinality of \mathcal{N} is \leq the cardinality of \mathcal{M} , and \mathcal{N} is elementarily equivalent to \mathcal{M} .

2611 Theorem. Let \mathcal{G} be a set of sentences consistent in $\mathcal{F}^=$. Then \mathcal{G} has a frugal equality-model.

2612 Extended Completeness and Soundness Theorem. Let \mathcal{S} be a set of sentences and \mathbf{A} be a wff of $\mathcal{F}^=$. Then $\mathcal{S} \vdash_{\mathcal{F}^=} \mathbf{A}$ iff \mathbf{A} is valid in all frugal equality-models of \mathcal{S} .

2613 Löwenheim-Skolem-Tarski Theorem. If \mathcal{S} is a set of sentences of $\mathcal{F}^=$ which has an infinite equality-model, and K is any cardinal number such that $K \geq \text{card}(\mathcal{L}(\mathcal{F}^=))$, then \mathcal{S} has an equality-model of cardinality K .

2614 Proposition. Suppose \mathcal{F} is a formulation of first-order logic such that $\mathcal{L}(\mathcal{F}) = \mathcal{L}(\mathcal{F}^=)$. For any wff \mathbf{D} of \mathcal{F} , let \mathbf{B} be the conjunction of all wffs of the form

$$\forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n [x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \mathbf{f} x_1 \dots x_n = \mathbf{f} y_1 \dots y_n],$$

where \mathbf{f} is a function symbol which occurs in \mathbf{D} , and let \mathbf{C} be the conjunction of all wffs of the form

$$\begin{aligned} \forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n [x_1 = y_1 \wedge \dots \wedge x_n = y_n \supset \\ \bullet \mathbf{P} x_1 \dots x_n \supset \mathbf{P} y_1 \dots y_n], \end{aligned}$$

where \mathbf{P} is $=$ or a predicate symbol which occurs in \mathbf{D} . Then $\vdash_{\mathcal{F}^=}\mathbf{D}$ iff $\vdash_{\mathcal{F}} \forall x[x = x] \wedge \mathbf{B} \wedge \mathbf{C} \supset \mathbf{D}$.

3100 Lemma. If $\vdash_{\mathcal{G}} \mathbf{M} \vee \sim\sim \mathbf{E}$ then $\vdash_{\mathcal{G}} \mathbf{M} \vee \mathbf{E}$.

3101 Theorem. For every wff \mathbf{A} of \mathcal{F} , $\vdash_{\mathcal{F}} \mathbf{A}$ iff $\vdash_{\mathcal{G}} \mathbf{A}$.

3102 Corollary. If $\vdash_{\mathcal{G}} \mathbf{M} \vee \mathbf{A}$ and $\vdash_{\mathcal{G}} \sim \mathbf{A} \vee \mathbf{N}$, then $\vdash_{\mathcal{G}} \mathbf{M} \vee \mathbf{N}$.

3200 Theorem. Let \mathcal{S} be a set of sentences of \mathcal{F} . \mathcal{S} has no model iff \mathcal{S} has a closed semantic tableau.

3300 Lemma. For any sentence \mathbf{B} of \mathcal{F} , $\vdash_{\mathcal{F}^*} (\star\mathbf{B}) \supset \mathbf{B}$.

3301 Theorem. Let \mathcal{G} be a set of sentences of \mathcal{F} and let $\mathcal{G}^* = \{\star\mathbf{B} \mid \mathbf{B} \in \mathcal{G}\}$. Then \mathcal{G}^* has a model iff \mathcal{G} has a model, and $\mathcal{F}^* \cup \mathcal{G}^*$ is a conservative extension $\mathcal{F} \cup \mathcal{G}$.

3302 Corollary. A sentence \mathbf{B} of \mathcal{F} is satisfiable iff $\star\mathbf{B}$ is satisfiable.

3303 Corollary. Let \mathbf{A} be any wff, and let $\overline{\mathbf{A}}$ be its universal closure. \mathbf{A} is valid iff $\star \sim \overline{\mathbf{A}}$ is unsatisfiable.

3400 Theorem. Let \mathcal{S} be any set of universal sentences. \mathcal{S} is \mathcal{R} -refutable iff \mathcal{S} is unsatisfiable.

3401 Theorem. Let \mathbf{A} be any wff, let $\overline{\mathbf{A}}$ be its universal closure, and let \mathbf{M} be the quantifier-free wff obtained by deleting all quantifiers from $\star \sim \overline{\mathbf{A}}$. Let $\mathbf{N}_1 \wedge \dots \wedge \mathbf{N}_k$ be a conjunctive normal form of \mathbf{M} (so each \mathbf{N}_i is a disjunction of literals).

The following conditions are equivalent:

- (a) $\models \mathbf{A}$.
- (b) $\{\mathbf{N}_1, \dots, \mathbf{N}_k\}$ has no model.
- (c) \square can be derived from $\{\mathbf{N}_1, \dots, \mathbf{N}_k\}$ by using rules $\mathcal{R}1$, $\mathcal{R}2$, $\mathcal{R}3$, and $\mathcal{R}4$.

3500 Lemma. Let \mathbf{D} be universal nnf, and let \mathbf{H} be a c-instance of \mathbf{D} . Then there is a conjunction \mathbf{M} of simple instances of \mathbf{D} such that $\mathbf{M} \equiv \mathbf{H}$ is tautologous; also, \mathbf{M} and \mathbf{H} contain exactly the same atomic wffs.

3501 Lemma. Let \mathbf{G} and \mathbf{H} be c-instances of a universal nnf \mathbf{D} . Then there is a c-instance \mathbf{K} of \mathbf{D} such that $\models \mathbf{K} \supset [\mathbf{G} \wedge \mathbf{H}]$.

3502 Lemma. Let \mathbf{D} be a universal nnf. If no c-instance of \mathbf{D} is a t-f contradiction, there is a truth assignment which verifies every c-instance of \mathbf{D} .

3503 Herbrand's Theorem. Let \mathbf{A} be a wff, let $\overline{\mathbf{A}}$ be its universal closure, and let \mathbf{D} be a negation normal form of the Skolemization $\star \sim \overline{\mathbf{A}}$ of $\sim \overline{\mathbf{A}}$. Then \mathbf{A} is valid if and only if \mathbf{D} has a compound Herbrand instance which is a truth-functional contradiction.

3504 Corollary. Let \mathbf{A} be a wff, let $\overline{\mathbf{A}}$ be its universal closure, and let \mathbf{D} be a negation normal form of the Skolemization $\star \sim \overline{\mathbf{A}}$ of $\sim \overline{\mathbf{A}}$. Then \mathbf{A} is valid if and only if some conjunction of simple Herbrand instances of \mathbf{D} is a truth-functional contradiction.

3505 Corollary. Let \mathbf{B} be a sentence, and let \mathbf{D} be a negation normal form of the Skolemization $\star \mathbf{B}$ of \mathbf{B} . Then \mathbf{B} is satisfiable if and only if the set of simple Herbrand instances of \mathbf{D} is t-f satisfiable.

3600 Proposition. If \mathcal{W} is a set of expressions and θ and τ are most general unifiers of \mathcal{W} , then $\theta \approx \tau$.

3601 Unification Theorem. There is an algorithm which can be applied to an arbitrary nonempty finite set \mathcal{W} of expressions to determine whether \mathcal{W} is unifiable. The algorithm always terminates, and produces a most general unifier of \mathcal{W} if \mathcal{W} is unifiable.

4000 Proposition. If \mathbf{B} and \mathbf{D} are duals of \mathbf{A} , then $\vdash \mathbf{B} \equiv \mathbf{D}$.

4001 Corollary. For any wff \mathbf{A} of \mathcal{F} , $\vdash \mathbf{A}^{dd} \equiv \mathbf{A}$.

4002 Theorem. Let $\mathcal{M} = \langle \mathcal{D}, \mathcal{J} \rangle$ be any interpretation. Define \mathcal{J}' to agree with \mathcal{J} on all individual and function constants, and for all n -ary predicate constants \mathbf{P} , let $(\mathcal{J}'\mathbf{P})d_1 \dots d_n = \sim (\mathcal{J}\mathbf{P})d_1 \dots d_n$ for all $d_1, \dots, d_n \in \mathcal{D}$. Let $\mathcal{M}' = \langle \mathcal{D}, \mathcal{J}' \rangle$. Then for any wff \mathbf{A} and assignment φ , $\mathcal{V}_\varphi^{\mathcal{M}} \mathbf{A} = \sim \mathcal{V}_\varphi^{\mathcal{M}'} \mathbf{A}$.

4003 Corollary. If $\vdash \mathbf{A}$, and \mathbf{B} is any dual of \mathbf{A} , then $\vdash \sim \mathbf{B}$.

4004 Corollary. If $\vdash \mathbf{A} \supset \mathbf{B}$ then $\vdash \mathbf{B}^d \supset \mathbf{A}^d$.

4005 Corollary. If $\vdash \mathbf{A} \equiv \mathbf{B}$, then $\vdash \mathbf{A}^d \equiv \mathbf{B}^d$.

4100 Craig-Lyndon Interpolation Theorem. If $\mathbf{H} \supset \mathbf{K}$ is a valid sentence of \mathcal{F}_f , then there is an interpolation sentence for $\mathbf{H} \supset \mathbf{K}$.

4200 Beth's Definability Theorem. Let \mathbf{H} be a sentence containing a k -ary predicate constant \mathbf{P} . Then \mathbf{H} defines \mathbf{P} implicitly iff \mathbf{P} is explicitly definable from \mathbf{H} .

5100 Principle of Induction on the Construction of a Wff. Let \mathcal{R} be a property of formulas, and let $\mathcal{R}(\mathbf{A})$ mean that \mathbf{A} has property \mathcal{R} . Suppose

- (1) Every formula consisting of a single variable or constant standing alone has property \mathcal{R} .
- (2) Whenever $\mathbf{A}_{\alpha\beta}$ is a wff $_{(\alpha)}$ and \mathbf{B}_β is a wff $_\beta$ and $\mathcal{R}(\mathbf{A}_{\alpha\beta})$ and $\mathcal{R}(\mathbf{B}_\beta)$, then $\mathcal{R}([\mathbf{A}_{\alpha\beta}\mathbf{B}_\beta])$.
- (3) Whenever \mathbf{x}_β is a variable and $\mathcal{R}(\mathbf{A}_\alpha)$, then $\mathcal{R}([\lambda\mathbf{x}_\beta\mathbf{A}_\alpha])$.

Then every wff of \mathcal{Q}_0 has property \mathcal{R} .

5200. $\vdash \mathbf{A}_\alpha = \mathbf{A}_\alpha$

5201 Equality Rules (= Rules).

If $\mathcal{H} \vdash \mathbf{A}$ and $\mathcal{H} \vdash \mathbf{A} \equiv \mathbf{B}$ then $\mathcal{H} \vdash \mathbf{B}$.

If $\mathcal{H} \vdash \mathbf{A}_\alpha = \mathbf{B}_\alpha$ then $\mathcal{H} \vdash \mathbf{B}_\alpha = \mathbf{A}_\alpha$.

If $\mathcal{H} \vdash \mathbf{A}_\alpha = \mathbf{B}_\alpha$ and $\mathcal{H} \vdash \mathbf{B}_\alpha = \mathbf{C}_\alpha$ then $\mathcal{H} \vdash \mathbf{A}_\alpha = \mathbf{C}_\alpha$.

If $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$ and $\mathcal{H} \vdash \mathbf{C}_\beta = \mathbf{D}_\beta$ then $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{D}_\beta$.

If $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta} = \mathbf{B}_{\alpha\beta}$ then $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{B}_{\alpha\beta}\mathbf{C}_\beta$.

If $\mathcal{H} \vdash \mathbf{C}_\beta = \mathbf{D}_\beta$ then $\mathcal{H} \vdash \mathbf{A}_{\alpha\beta}\mathbf{C}_\beta = \mathbf{A}_{\alpha\beta}\mathbf{D}_\beta$.

5202 Rule RR. If $\vdash \mathbf{A}_\alpha = \mathbf{B}_\alpha$ or $\vdash \mathbf{B}_\alpha = \mathbf{A}_\alpha$, and if $\mathcal{H} \vdash \mathbf{C}$, then $\mathcal{H} \vdash \mathbf{D}$, where \mathbf{D} is obtained from \mathbf{C} by replacing one occurrence of \mathbf{A}_α in \mathbf{C} by an occurrence of \mathbf{B}_α , and the occurrence of \mathbf{A}_α is not an occurrence of a variable immediately preceded by λ .

5203. $\vdash [\lambda\mathbf{x}_\alpha\mathbf{B}_\beta] \mathbf{A}_\alpha = S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} \mathbf{B}_\beta$ provided no variable in \mathbf{A}_α is bound in \mathbf{B}_β .

5204. If $\vdash \mathbf{B}_\beta = \mathbf{C}_\beta$, then $\vdash S_{\mathbf{A}_\alpha}^{\mathbf{x}_\alpha} [\mathbf{B}_\beta = \mathbf{C}_\beta]$, provided no variable in \mathbf{A}_α is bound in \mathbf{B}_β or \mathbf{C}_β .

5205. $\vdash f_{\alpha\beta} = [\lambda\mathbf{y}_\beta \bullet f_{\alpha\beta}\mathbf{y}_\beta]$

5206 $\vdash [\lambda\mathbf{x}_\beta\mathbf{A}_\alpha] = [\lambda\mathbf{z}_\beta S_{\mathbf{z}_\beta}^{\mathbf{x}_\beta} \mathbf{A}_\alpha]$, provided that \mathbf{z}_β does not occur free in \mathbf{A}_α and \mathbf{z}_β is free for \mathbf{x}_β in \mathbf{A}_α .

5207 $\vdash [\lambda x_\alpha B_\beta] A_\alpha = S_{A_\alpha}^{x_\alpha} B_\beta$, provided that A_α is free for x_α in B_β .

5208 $\vdash [\lambda x_{\alpha_1}^1 \dots \lambda x_{\alpha_n}^n B_\beta] x_{\alpha_1}^1 \dots x_{\alpha_n}^n = B_\beta$ where $n \geq 1$.

5209. If $\vdash B_\beta = C_\beta$, then $\vdash S_{A_\alpha}^{x_\alpha} [B_\beta = C_\beta]$, provided A_α is free for x_α in $[B_\beta = C_\beta]$.

5210 $\vdash T = [B_\beta = B_\beta]$

5211. $\vdash [T \wedge T] = T$

5212. $\vdash T \wedge T$

5213. If $\vdash A_\alpha = B_\alpha$ and $\vdash C_\beta = D_\beta$ then $\vdash [A_\alpha = B_\alpha] \wedge [C_\beta = D_\beta]$.

5214. $\vdash T \wedge F = F$

5215 Universal Instantiation (VI). If $\mathcal{H} \vdash \forall x_\alpha B$, then $\mathcal{H} \vdash S_{A_\alpha}^{x_\alpha} B$, provided that A_α is free for x_α in B .

5216. $\vdash [T \wedge A] = A$

5217. $\vdash [T = F] = F$

5218 $\vdash [T = A] = A$

5219 Rule T. $\mathcal{H} \vdash A$ if and only if $\mathcal{H} \vdash T = A$; $\mathcal{H} \vdash A$ if and only if $\mathcal{H} \vdash A = T$.

5220 Rule of Universal Generalization (Gen). If $\mathcal{H} \vdash A$ then $\mathcal{H} \vdash \forall x_\alpha A$, provided that x_α is not free in any wff in \mathcal{H} .

5221_n Rule of Substitution (Sub). If $\mathcal{H} \vdash B$, then $\mathcal{H} \vdash S_{A_{\alpha_1}^1 \dots A_{\alpha_n}^n}^{x_{\alpha_1}^1 \dots x_{\alpha_n}^n} B$, provided that the variables x^1, \dots, x^n are distinct from one another and from all free variables of wffs in \mathcal{H} , and that $A_{\alpha_i}^i$ is free for $x_{\alpha_i}^i$ in B for all i ($1 \leq i \leq n$).

5222 Rule of Cases. If $\mathcal{H} \vdash S_T^{x_o} A$ and $\mathcal{H} \vdash S_F^{x_o} A$, then $\mathcal{H} \vdash A$.

5223. $\vdash [T \supset y_o] = y_o$

5224 Modus Ponens (MP). If $\mathcal{H} \vdash A$ and $\mathcal{H} \vdash A \supset B$, then $\mathcal{H} \vdash B$.

5225. $\vdash \Pi_{o(o\alpha)} f_{o\alpha} \supset f_{o\alpha} x_\alpha$

5226. $\vdash \forall x_\alpha B \supset S_{A_\alpha}^{x_\alpha} B$ provided that A_α is free for x_α in B .

5227. $\vdash F \supset x_o$

5228. $\vdash [T \supset T] = T$; $\vdash [T \supset F] = F$; $\vdash [F \supset T] = T$; $\vdash [F \supset F] = T$

5229. $\vdash [T \wedge T] = T$; $\vdash [T \wedge F] = F$; $\vdash [F \wedge T] = F$; $\vdash [F \wedge F] = F$

5230. $\vdash [T = T] = T$; $\vdash [T = F] = F$; $\vdash [F = T] = F$; $\vdash [F = F] = T$

5231. $\vdash \sim T = F$; $\vdash \sim F = T$

5232. $\vdash T \vee T = T$; $\vdash T \vee F = T$; $\vdash F \vee T = T$; $\vdash F \vee F = F$

5233. If A is a tautology, then $\vdash A$.

5234 Rule P. If $\mathcal{H} \vdash A^1, \dots, \mathcal{H} \vdash A^n$, and if $[A^1 \wedge \dots \wedge A^n] \supset B$ is tautologous, then $\mathcal{H} \vdash B$. Also, if B is tautologous, then $\mathcal{H} \vdash B$.

5235. $\vdash \forall x_\alpha [A \vee B] \supset \neg A \vee \forall x_\alpha B$ provided x_α is not free in A .

5236 “Rule Q”. If $\mathcal{H} \vdash A^1, \dots$, and $\mathcal{H} \vdash A^n$ (where $n \geq 0$), and if \bar{A}^i can be obtained from A^i by generalizing on zero or more free variables of A^i which do not occur free in \mathcal{H} , and if $[\bar{A}^1 \wedge \dots \wedge \bar{A}^n \supset B]$ is a substitution instance of a theorem of quantification theory, then $\mathcal{H} \vdash B$.

5237 $\supset \forall$ Rule. If $\mathcal{H} \vdash A \supset B$, then $\mathcal{H} \vdash A \supset \forall x B$, provided that x_α is not free in A or in \mathcal{H} .

5238. $\vdash [\lambda x_{\beta_1}^1 \dots \lambda x_{\beta_n}^n A_\alpha] = [\lambda x_{\beta_1}^1 \dots \lambda x_{\beta_n}^n B_\alpha] = \forall x_{\beta_1}^1 \dots \forall x_{\beta_n}^n [A_\alpha = B_\alpha]$

5239. $\vdash \forall x_{\beta_1}^1 \dots \forall x_{\beta_n}^n [A_\alpha = B_\alpha] \supset \neg C = D$, where $n \geq 0$ and

- (a) D is obtained from C and $[A_\alpha = B_\alpha]$ as in Rule R;
- (b) $x_{\beta_1}^1, \dots, x_{\beta_n}^n$ is a complete list of those variables x_β such that x_β occurs free in $[A_\alpha = B_\alpha]$, and the occurrence of A_α in C (which is replaced by B_α in D) is in a wf part of C of the form $[\lambda x_\beta E_\gamma]$.

5240 Deduction Theorem. If $\mathcal{H}, H \vdash P$ then $\mathcal{H} \vdash H \supset P$.

5241. If $\mathcal{H} \vdash A$ and $\mathcal{H} \subseteq \mathcal{G}$ then $\mathcal{G} \vdash A$.

5242 Rule of Existential Generalization (\exists Gen). If $\mathcal{H} \vdash S_{A_\alpha}^{x_\alpha} B$, and A_α is free for x_α in B , then $\mathcal{H} \vdash \exists x_\alpha B$.

5243 Comprehension Theorem.

$\vdash \exists u_{\beta\alpha_n\dots\alpha_1} \forall x_{\alpha_1}^1 \dots \forall x_{\alpha_n}^n \bullet u_{\beta\alpha_n\dots\alpha_1} x_{\alpha_1}^1 \dots x_{\alpha_n}^n = B_\beta$
when $n \geq 0$ and $u_{\beta\alpha_n\dots\alpha_1}$ is not free in B_β .

5244 Existential Rule (\exists Rule). If $\mathcal{H}, B \vdash A$, and x_α is not free in \mathcal{H} or in A , then $\mathcal{H}, \exists x_\alpha B \vdash A$.

5245 Rule C. If $\mathcal{H} \vdash \exists x_\alpha B$ and $\mathcal{H}, S_{y_\alpha}^{x_\alpha} B \vdash A$, where y_α is free for x_α in B and y_α is not free in \mathcal{H} , $\exists x_\alpha B$, or A , then $\mathcal{H} \vdash A$.

5300. $\vdash [x_\alpha = y_\alpha] \supset \bullet h_{\beta\alpha} x_\alpha = h_{\beta\alpha} y_\alpha$

5301. $\vdash [x_\alpha = y_\alpha] \wedge [f_{\beta\alpha} = g_{\beta\alpha}] \supset \bullet f_{\beta\alpha} x_\alpha = g_{\beta\alpha} y_\alpha$

5302. $\vdash [x_\alpha = y_\alpha] = [y_\alpha = x_\alpha]$

5303. $\vdash [x_\alpha = y_\alpha] \supset \bullet [x_\alpha = z_\alpha] = [y_\alpha = z_\alpha]$

5304. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \exists y_\alpha \bullet p_{o\alpha} = Q_{o\alpha\alpha} y_\alpha$

5305. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \exists y_\alpha \forall z_\alpha \bullet p_{o\alpha} z_\alpha = \bullet y_\alpha = z_\alpha$

5306. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \exists y_\alpha \bullet p_{o\alpha} y_\alpha \wedge \forall z_\alpha \bullet p_{o\alpha} z_\alpha \supset y_\alpha = z$

5307. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha = \bullet \exists y_\alpha p_{o\alpha} y_\alpha \wedge \forall y_\alpha \forall z_\alpha \bullet p_{o\alpha} y_\alpha \wedge p_{o\alpha} z_\alpha \supset y_\alpha = z_\alpha$

5308. $\vdash \iota_{o(oo)} [Q_{ooo} y_o] = y_o$

5309 $^\gamma$. $\vdash \iota_{\gamma(o\gamma)} [Q_{o\gamma\gamma} y_\gamma] = y_\gamma$ for each type symbol γ .

5310. $\vdash \forall z_\alpha [p_{o\alpha} z_\alpha = \bullet y_\alpha = z_\alpha] \supset \bullet \iota_{\alpha(o\alpha)} p_{o\alpha} = y_\alpha$

5311. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha \supset p_{o\alpha} [\iota_{\alpha(o\alpha)} p_{o\alpha}]$

5312. $\vdash \exists_1 y_\alpha p_{o\alpha} y_\alpha \supset \forall z_\alpha \bullet p_{o\alpha} z_\alpha = \bullet \iota_{\alpha(o\alpha)} p_{o\alpha} = z_\alpha$

5313. $\vdash [C_{\gamma o\gamma\gamma} x_\gamma y_\gamma T_o = x_\gamma] \wedge \bullet C_{\gamma o\gamma\gamma} x_\gamma y_\gamma F_o = y_\gamma$

5400 Proposition. Let \mathcal{M} be a general model, \mathbf{A}_α a wff, and φ and ψ assignments which agree on all free variables of \mathbf{A}_α . Then $\mathcal{V}_\varphi^\mathcal{M} \mathbf{A}_\alpha = \mathcal{V}_\psi^\mathcal{M} \mathbf{A}_\alpha$.

5401 Lemma. Let \mathcal{M} be a general model, and φ be an assignment into \mathcal{M} .

- (a) $\mathcal{V}_\varphi^\mathcal{M} [[\lambda \mathbf{x}_\alpha \mathbf{B}_\beta] \mathbf{A}_\alpha] = \mathcal{V}_{(\varphi: \mathbf{x}_\alpha / \mathcal{V}_\varphi^\mathcal{M} \mathbf{A}_\alpha)}^\mathcal{M} \mathbf{B}_\alpha$.
- (b) $\mathcal{V}_\varphi^\mathcal{M} [\mathbf{A}_\alpha = \mathbf{B}_\alpha] = \top$ iff $\mathcal{V}_\varphi^\mathcal{M} \mathbf{A}_\alpha = \mathcal{V}_\varphi^\mathcal{M} \mathbf{B}_\alpha$.
- (c) $\mathcal{V}_\varphi^\mathcal{M} T_o = \top$.
- (d) $\mathcal{V}_\varphi^\mathcal{M} F_o = \mathbb{F}$.
- (e) If $x, y \in \mathcal{D}_o$, then $(\mathcal{V}^\mathcal{M} \wedge_{ooo}) xy = \begin{cases} \top & \text{if } x = \top \text{ and } y = \top; \\ \mathbb{F} & \text{otherwise.} \end{cases}$
- (f) If $x, y \in \mathcal{D}_o$, then $(\mathcal{V}^\mathcal{M} \supset_{ooo}) xy = \begin{cases} \top & \text{if } x = \mathbb{F} \text{ or } y = \top; \\ \mathbb{F} & \text{if } x = \top \text{ and } y = \mathbb{F}. \end{cases}$
- (g) $\mathcal{M} \models_\varphi \forall \mathbf{x} \mathbf{A}$ iff $\mathcal{M} \models_\psi \mathbf{A}$ for all assignments ψ which agree with φ off \mathbf{x}_α .

5402 Soundness Theorem.

- (a) Every theorem of \mathcal{Q}_0 is valid in the general sense, and hence in the standard sense.
- (b) If \mathcal{G} is a set of wffs_o and \mathcal{M} is a model for \mathcal{G} and $\mathcal{G} \vdash \mathbf{A}$, then $\mathcal{M} \models \mathbf{A}$.

5403 Consistency Theorem \mathcal{Q}_0 is consistent, and every set of wffs_o which has a model is consistent.

5404 Theorem. Every finite model for \mathcal{Q}_0 is standard.

5405 Proposition. An isomorphism of general models is a one-to-one mapping.

5406 Proposition. Let \mathcal{N} and \mathcal{M} be general models for \mathcal{Q}_0 , and let τ be an isomorphism from \mathcal{N} to \mathcal{M} . For each assignment φ into \mathcal{N} , let $\tau \circ \varphi$ be the assignment into \mathcal{M} such that for each variable \mathbf{x}_γ , $(\tau \circ \varphi)\mathbf{x}_\gamma = \tau(\varphi \mathbf{x}_\gamma)$. Then for every wff \mathbf{C}_γ of \mathcal{Q}_0 , $\tau(\mathcal{V}_\varphi^\mathcal{N} \mathbf{C}_\gamma) = \mathcal{V}_{\tau \circ \varphi}^\mathcal{M} \mathbf{C}_\gamma$.

5407 Corollary. If \mathcal{N} and \mathcal{M} are isomorphic general models of \mathcal{Q}_0 , then $\mathcal{M} \models \mathbf{C}$ iff $\mathcal{N} \models \mathbf{C}$.

5500 Extension Lemma. Let \mathcal{G} be any consistent set of sentences of \mathcal{Q}_0 . Then there is an expansion \mathcal{Q}_0^+ of \mathcal{Q}_0 and a set of \mathcal{H} of sentences of \mathcal{Q}_0^+ such that

- (1) $\mathcal{G} \subseteq \mathcal{H}$.
- (2) \mathcal{H} is consistent.
- (3) \mathcal{H} is complete in \mathcal{Q}_0^+ .
- (4) \mathcal{H} is extensionally complete in \mathcal{Q}_0^+ .
- (5) $\text{card}(\mathcal{L}(\mathcal{Q}_0^+)) = \text{card}(\mathcal{L}(\mathcal{Q}))$.

5501 Henkin's Theorem. Every consistent set of sentences of \mathcal{Q}_0 has a frugal general model.

5502 Henkin's Completeness and Soundness Theorem. Let \mathbf{A} be a wff_o and let \mathcal{G} be a set of sentences.

- (a) $\vdash \mathbf{A}$ iff $\models \mathbf{A}$.
- (b) $\mathcal{G} \vdash \mathbf{A}$ iff \mathbf{A} is valid in every general model for \mathcal{G} .
- (c) $\mathcal{G} \vdash \mathbf{A}$ iff \mathbf{A} is valid in every frugal general model for \mathcal{G} .

5503 Weak Compactness Theorem. If every finite subset of the set \mathcal{S} of sentences has a general model, then \mathcal{S} has a general model.

5504 Lemma. Let $\mathcal{M} = \langle \{\mathcal{D}_\alpha\}_\alpha, \mathcal{J} \rangle$ be a frugal model for \mathcal{Q}_0 such that $\text{card}(\mathcal{D}_i) = \text{card}(\mathcal{L}(\mathcal{Q}_0))$. Then \mathcal{M} is nonstandard.

5505 Corollary. Every infinite frugal model for a countable formulation of \mathcal{Q}_0 is nonstandard.

5506 Theorem. Let \mathcal{H} be any set of sentences of \mathcal{Q}_0 which has infinite models. Then \mathcal{H} has nonstandard models.

6000. $\vdash \sim (\forall \mathbf{x}_\alpha \in \mathbf{M}_{o\alpha}) \mathbf{A} = (\exists \mathbf{x}_\alpha \in \mathbf{M}_{o\alpha}) \sim \mathbf{A}$ provided \mathbf{x}_α is not free in $\mathbf{M}_{o\alpha}$.

6001. $\vdash \sim (\exists \mathbf{x}_\alpha \in \mathbf{M}_{o\alpha}) \mathbf{A} = (\forall \mathbf{x}_\alpha \in \mathbf{M}_{o\alpha}) \sim \mathbf{A}$ provided \mathbf{x}_α is not free in $\mathbf{M}_{o\alpha}$.

6100. $\vdash \mathbb{N}_{o\sigma} 0_\sigma$ by the def of \mathbb{N}

6101. $\vdash \forall x_\sigma \bullet \mathbb{N}_{o\sigma} x_\sigma \supset \mathbb{N}_{o\sigma} \bullet S_{o\sigma} x_\sigma$

6102. The Induction Theorem.

$$\vdash \forall p_{o\sigma} \bullet [p_{o\sigma} 0_\sigma \wedge \dot{\forall} x_\sigma \bullet p_{o\sigma} x_\sigma \supset p_{o\sigma} \bullet S_{o\sigma} x_\sigma] \supset \dot{\forall} x_\sigma p_{o\sigma} x_\sigma$$

6103. $\vdash p_{o\sigma} 0_\sigma \wedge \dot{\forall} x_\sigma p_{o\sigma} S_{o\sigma} x_\sigma \supset \dot{\forall} n_\sigma p_{o\sigma} n_\sigma$

6104. $\vdash \dot{\forall} n_\sigma \bullet n_\sigma = 0 \vee \dot{\exists} m_\sigma \bullet n_\sigma = S_{\sigma\sigma}m_\sigma$

6105. $\vdash \forall n_\sigma \bullet S_{\sigma\sigma}n_\sigma \neq 0_\sigma$

6106. $\vdash \dot{\forall} n_\sigma \bullet S_{\sigma\sigma}n_\sigma \neq 0_\sigma$

6107. $\vdash \dot{\forall} n_\sigma \forall p_{o1} \bullet S_{n_\sigma}p_{o1} \wedge p_{o1}w_i \supset n_\sigma [\lambda t_i \bullet t_i \neq w_i \wedge p_{o1}t_i]$

6108. $\vdash \dot{\forall} n_\sigma \bullet n_\sigma p_{o1} \supset \exists w_i \sim p_{o1}w_i$

6109. $\vdash \dot{\forall} n_\sigma \dot{\forall} m_\sigma \bullet S_{\sigma\sigma}n_\sigma = S_{\sigma\sigma}m_\sigma \supset n_\sigma = m_\sigma$

6110 $\vdash \mathbb{N}\bar{n}$ where n is any natural number.

6111. If k is a natural number greater than 0, $\vdash \dot{\forall} x_\sigma \bullet S^k x_\sigma \neq x_\sigma$.

6112 Theorem. $Q_{o\sigma\sigma}$ represents the equality relation. Thus, if m and n are distinct natural numbers, $\vdash \bar{m} \neq \bar{n}$.

6113 Theorem. Let r be any relation between natural numbers. Then r is representable iff χ_r is representable.

6200. $\vdash \dot{\forall} y_\sigma \bullet 0 \leq y_\sigma$

6201. $\vdash x_\sigma \leq x_\sigma$

6202. $\vdash x_\sigma = y_\sigma \supset x_\sigma \leq y_\sigma$

6203. $\vdash x_\sigma \leq Sx_\sigma$

6204. $\vdash x_\sigma \leq y_\sigma \wedge y_\sigma \leq z_\sigma \supset x_\sigma \leq z_\sigma$

6205. $\vdash \dot{\forall} x_\sigma \bullet x_\sigma \leq y_\sigma \supset \bullet x_\sigma = y_\sigma \vee \dot{\exists} z_\sigma \bullet x_\sigma \leq z_\sigma \wedge Sz = y$

6206. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \bullet x_\sigma \leq Sy_\sigma \equiv \bullet x_\sigma \leq y_\sigma \vee x_\sigma = Sy_\sigma$

6207⁰. $\vdash \dot{\forall} x_\sigma \bullet x \leq 0 \equiv \bullet x = 0$

6207¹. $\vdash \dot{\forall} x_\sigma \bullet x \leq \bar{1} \equiv \bullet x = 0 \vee x = \bar{1}$

6207^k. $\vdash \dot{\forall} x_\sigma \bullet x \leq \bar{k} \equiv \bullet x = 0 \vee x = \bar{1} \vee \dots \vee x = \bar{k}$

6208 Theorem. $\leq_{o\sigma\sigma}$ represents the relation \leq between natural numbers.

6209^k. $\vdash \dot{\forall} y_\sigma [y_\sigma \leq \bar{k} \supset p_{o\sigma} y_\sigma] = \blacksquare p_{o\sigma} 0 \wedge \dots \wedge p_{o\sigma} \bar{k}$

6210^k. $\vdash \dot{\exists} y_\sigma [y_\sigma \leq \bar{k} \wedge p_{o\sigma} y_\sigma] = \blacksquare p_{o\sigma} 0 \vee \dots \vee p_{o\sigma} \bar{k}$

6211. $\vdash \dot{\forall} y_\sigma \dot{\forall} x_\sigma \leq y_\sigma \supset Sx_\sigma \leq Sy_\sigma$

6212. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \blacksquare x \leq y \vee y \leq x$

6213. $\vdash \dot{\forall} y_\sigma \dot{\forall} x_\sigma \blacksquare x \leq y \wedge y \leq x \supset x = y$

6214. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \blacksquare \sim x \leq y \supset Sy \leq x$

6215. $\vdash Infin^\sigma$

6300. $\vdash N_{o\sigma} t_\sigma \wedge p_{o\sigma} t_\sigma \supset \blacksquare N_{o\sigma} [\mu_{\sigma(o\sigma)} p_{o\sigma}] \wedge p_{o\sigma} [\mu_{\sigma(o\sigma)} p_{o\sigma}] \wedge \mu_{\sigma(o\sigma)} p_{o\sigma} \leq t_\sigma$

6301. $\vdash N_{o\sigma} t_\sigma \wedge p_{o\sigma} t_\sigma \wedge \dot{\forall} y_\sigma [y_\sigma \leq t_\sigma \supset \blacksquare p_{o\sigma} y_\sigma \supset y_\sigma = t_\sigma] \supset \mu_{\sigma(o\sigma)} p_{o\sigma} = t_\sigma$

6302 Theorem. If r is a representable $(k+1)$ -ary regular numerical relation, then $((\lambda n_1, \dots, n_k)(\mu m)r(n_1, \dots, n_k, m))$ is a representable function.

6303 Theorem. If g is a representable $(k+1)$ -ary regular numerical function, and h is obtained from g by minimization, then h is representable.

6400. $\vdash [R_{\sigma\sigma\sigma(\sigma\sigma\sigma)} h_{\sigma\sigma\sigma} g_\sigma] 0_\sigma = g_\sigma \wedge \dot{\forall} n_\sigma \blacksquare [Rh_{\sigma\sigma\sigma} g_\sigma][S_{\sigma\sigma} n] = h_{\sigma\sigma\sigma} n_\sigma \blacksquare [Rh_{\sigma\sigma\sigma} g_\sigma] n_\sigma$

6401 Theorem. Every general recursive function is representable in \mathcal{Q}_0^∞ .

6500 Constant Functions. For all $k \geq 0$ and $n \geq 0$ $((\lambda \vec{x}_k) n) \in \mathcal{PRF}^k$.

6501 Explicit Definitions. If $\mathcal{S} \subseteq \mathcal{PRF}$ and f is a numerical function explicitly definable from \mathcal{S} , then $f \in \mathcal{PRF}$.

6502 Generalized Primitive Recursion. Let $\mathcal{S} \subseteq \mathcal{PRF}$, and let g and h be k -ary and $(k+2)$ -ary functions, respectively, which are explicitly definable from \mathcal{S} . If for all natural numbers x_1, \dots, x_k , and y ,

$$f(x_1, \dots, x_i, 0, x_{i+1}, \dots, x_k) = g(\vec{x}_k), \text{ and}$$

$$f(x_1, \dots, x_i, Sy, x_{i+1}, \dots, x_k) = h(\vec{x}_k, f(x_1, \dots, x_i, y, x_{i+1}, \dots, x_k)),$$

then $f \in \mathcal{PR}$.

6503 + (addition) $\in \mathcal{PRF}^2$.

6504. If $f \in \mathcal{PRF}^{k+1}$ (where $k \geq 0$), then $((\lambda \vec{x}_k, y) \sum_{z \leq y} f(\vec{x}_k, z)) \in \mathcal{PRF}^{k+1}$.

6505. \times (multiplication) $\in \mathcal{PRF}^2$.

6506. If $f \in \mathcal{PRF}^{k+1}$ (where $k \geq 0$), then $((\lambda \vec{x}_k, y) \Pi_{z \leq y} f(\vec{x}_k, z)) \in \mathcal{PRF}^{k+1}$.

6507. $((\lambda x, y) x^y)$ (exponentiation) $\in \mathcal{PRF}^2$.

6508. ! (factorial) $\in \mathcal{PRF}^1$.

6509. δ (predecessor) $\in \mathcal{PRF}^1$.

6510. $-$ (modified subtraction or monus) $\in \mathcal{PRF}^2$.

6511. σ (signum) $\in \mathcal{PRF}^1$.

6512. $\bar{\sigma}$ (antisignum) $\in \mathcal{PRF}^1$.

6513. $<$ $\in \mathcal{PRR}^2$.

6514. \leq $\in \mathcal{PRR}^2$.

6515. $=$ $\in \mathcal{PRR}^2$

6516. If $\mathcal{S} \subseteq \mathcal{PR}$, and the numerical relation r is explicitly definable from \mathcal{S} using propositional connectives, then $r \in \mathcal{PR}$.

6517 Bounded Quantification. If $f \in \mathcal{PRF}^k$ (where $k \geq 0$) and $r \in \mathcal{PRR}^{k+1}$, then the following are in \mathcal{PRR}^k :

- (a) $(\lambda \vec{x}_k)(\exists z \leq f(\vec{x}_k))r(\vec{x}_k, z)$
- (b) $(\lambda \vec{x}_k)(\exists z < f(\vec{x}_k))r(\vec{x}_k, z)$
- (c) $(\lambda \vec{x}_k)(\forall z \leq f(\vec{x}_k))r(\vec{x}_k, z)$
- (d) $(\lambda \vec{x}_k)(\forall z < f(\vec{x}_k))r(\vec{x}_k, z)$

6518 Bounded Minimization and Maximization. If $f \in \mathcal{PRF}^k$ and $r \in \mathcal{PRR}^{k+1}$ (where $k \geq 0$), then the following are in \mathcal{PRF}^k :

- (a) $(\lambda \vec{x}_k)(\mu z \leq f(\vec{x}_k))r(\vec{x}_k, z)$

$$(b) (\lambda \vec{x}_k)(\text{Max } z \leq f(\vec{x}_k))r(\vec{x}_k, z)$$

6519 Definition by Cases. Let $g_1, \dots, g_n \in \mathcal{PRF}^k$ and $r_1, \dots, r_n \in \mathcal{PRR}^k$, and suppose the relations r_1, \dots, r_n are mutually exclusive and exhaustive, i.e., for all $\vec{x}_k \in \mathbb{N}^k$, there is exactly one r_i such that $r_i(\vec{x}_k)$. Let $f(\vec{x}_k) = g_i(\vec{x}_k)$ iff $r_i(\vec{x}_k)$ for $1 \leq i \leq n$. Then $f \in \mathcal{PRF}^k$.

6520. $((\lambda x, y)x \text{ Divides } y) \in \mathcal{PRR}^2$

6521. Let $\text{Prime}(x)$ iff x is a prime number. Then $\text{Prime} \in \mathcal{PRR}^1$.

6522. Let $\text{Pr}(n)$ be the nth odd prime number, where $\text{Pr}(0) = 2$. Then $\text{Pr} \in \mathcal{PRF}^1$.

6523. Let $n \mathcal{G} \ell x$ be the exponent of $\text{Pr}(n)$ in the prime factorization of x , with $n \mathcal{G} \ell 0 = 0$. Then $\mathcal{G} \ell \in \mathcal{PRF}^2$.

6524. Let $E(x)$ (the extent of the sequence represented by x) be the largest number n such that $\text{Pr}(n)$ divides x , with $E(0) = 0 = E(1)$. Then $E \in \mathcal{PRF}^1$.

6525. Let $x * y = x \times \prod_{i \leq E(y)} \text{Pr}(E(x) + 1 + i)^{i \mathcal{G} \ell y}$ if $1 < x$
 $= y$ if $x \leq 1$.

Then $* \in \mathcal{PRF}^2$.

6526. Let $Lg = (\lambda x)(\mu y \leq x)(\exists z \leq x)[x = z^y \wedge \text{Prime}(z)]$. Then $Lg \in \mathcal{PRF}^1$.

7000 Theorem. Let $\text{TypeSymbol} = ((\lambda n)n \text{ is the number of some type symbol})$. Then $\text{TypeSymbol} \in \mathcal{PRR}^1$.

7001 Theorem. Let $Q = (\lambda n)19 \exp((3 \exp(3^3 \times 5^n)) \times 5^n)$. Then $Q \in \mathcal{PRF}^1$, and for each type symbol α , $Q(" \alpha ") = 'Q_{o\alpha\alpha}'$.

7002 Theorem. Let $\text{Variable} = ((\lambda n)n \text{ is the number of a variable of } Q_0)$. Then $\text{Variable} \in \mathcal{PRR}^1$.

7003 Theorem. Let $\text{ProperSymbol} = ((\lambda n)n \text{ is the number of a variable or constant of } Q_0)$. Then $\text{ProperSymbol} \in \mathcal{PRR}^1$.

7004 Theorem. Let $\text{Wff} = ((\lambda x) \text{ there is a wff } \mathbf{A}_\alpha \text{ of } \mathcal{Q}_0 \text{ such that } x = \text{"A}_\alpha\text{"})$.

$$\begin{aligned}\text{Let Type}(x) &= \text{"}\alpha\text{" if } x = \text{"A}_\alpha\text{"}; \\ &= 0 \text{ if not Wff}(x).\end{aligned}$$

Then $\text{Wff} \in \mathcal{PRR}^1$ and $\text{Type} \in \mathcal{PRF}^1$.

7005 Theorem. Let $\text{Axiom} = ((\lambda n) \text{ there is an axiom } \mathbf{A} \text{ of } \mathcal{Q}_0^\infty \text{ such that } n = \text{"A"})$. Then $\text{Axiom} \in \mathcal{PRR}^1$.

7006 Theorem Let $\text{Rule} = (\lambda m, n, p)$ [there exist wfds \mathbf{A}_α , \mathbf{B}_α , \mathbf{C}_o , and \mathbf{D}_o such that $m = [\mathbf{A}_\alpha = \mathbf{B}_\alpha]$ and $n = \mathbf{C}_o$ and $p = \mathbf{D}_o$, and \mathbf{D}_o is obtained from $[\mathbf{A}_\alpha = \mathbf{B}_\alpha]$ and \mathbf{C}_o by Rule R]. Then $\text{Rule} \in \mathcal{PRR}^3$.

7007 Theorem. Let \mathcal{A} be any [primitive] recursively axiomatized pure extension of \mathcal{Q}_0 . Let $\text{Proof}^{\mathcal{A}} = (\lambda m, n)$ [There is a proof $\mathbf{P}^0, \dots, \mathbf{P}^k$ in \mathcal{A} such that $m = \text{"P}^k\text{"}$ and $n = \text{"P}^0, \dots, \text{P}^k\text{"}$]. Then $\text{Proof}^{\mathcal{A}}$ is a [primitive] recursive relation.

7008 Theorem. Let $\text{Num} = ((\lambda n) \text{"}\bar{n}\text{"})$. Then $\text{Num} \in \mathcal{PRF}^1$.

7009 Theorem. If \mathcal{A} is a consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ , then every numerical function and every numerical relation representable in \mathcal{A} is recursive.

7010 Theorem. If \mathcal{A} is any consistent recursively axiomatized pure extension of \mathcal{Q}_0^∞ , then a numerical function is representable in \mathcal{A} if and only if it is recursive.

7011 Normal Form Theorem. For each natural number $n \geq 1$ there is a relation $S_n \in \mathcal{PRR}^{n+2}$ such that for each n -ary recursive function f , there is a number j such that $(\lambda \vec{x}_n, y)S_n(j, \vec{x}_n, y)$ is regular, and

$$f = (\lambda \vec{x}_n)0 \text{ } \mathcal{G} \ell ((\mu y)S_n(j, \vec{x}_n, y)).$$

7012 Theorem. Every recursive function can be defined with at most one application of minimization.

7100 Proposition. If \mathcal{A} is an ω -consistent extension of \mathcal{Q}_0 , then \mathcal{A} is consistent.

7101 Gödel's Incompleteness Theorem [Gödel, 1931]. Let \mathcal{A} be any recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Let $\star_{o\sigma}$ be the wff

$$[\lambda x_\sigma \dot{\forall} y_\sigma \sim Proof_{o\sigma}^{\mathcal{A}} [\overline{128} * x_\sigma * [Num_{o\sigma} x_\sigma] * \overline{2048}] y_\sigma].$$

Let $m = “\star_{o\sigma}”$ and let $\star\star_o$ be $[\star_{o\sigma} \overline{m}]$.

- (1) If \mathcal{A} is consistent, then not $\vdash_{\mathcal{A}} \star\star$.
- (2) If \mathcal{A} is ω -consistent, then not $\vdash_{\mathcal{A}} \sim \star\star$.

7102 Corollary. Every consistent recursively axiomatized pure extension \mathcal{A} of \mathcal{Q}_0^∞ is ω -incomplete.

7103 Lemma. If $\vdash \mathbf{A}$, then $\vdash Theorem_{o\sigma} \overline{“A”}$.

7104 Lemma. $\vdash \dot{\forall} x_\sigma \blacksquare Theorem_{o\sigma} x_\sigma \supset Theorem_{o\sigma} [\overline{“[Theorem_{o\sigma} “A”] * Num x_\sigma * \overline{2048}”}]$

7105 Corollary. For any wff_o \mathbf{A} ,

$$\vdash Theorem_{o\sigma} \overline{“A”} \supset Theorem_{o\sigma} \overline{“[Theorem_{o\sigma} \overline{“A”}]”}.$$

7106 Lemma. $\vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \blacksquare Theorem_{o\sigma} [\overline{“[[Q_{ooo}] * x_\sigma * \overline{2048} * y_\sigma * \overline{2048}]”}] \supset \blacksquare Theorem_{o\sigma} x_\sigma \supset \blacksquare Theorem_{o\sigma} y_\sigma$.

7107 Corollary.

$$\vdash Theorem_{o\sigma} \overline{“A_o = B_o”} \supset \blacksquare Theorem_{o\sigma} \overline{“A_o”} \supset Theorem_{o\sigma} \overline{“B_o”}.$$

7108 Corollary. If $\vdash \mathbf{A} = \mathbf{B}$, then $\vdash Theorem_{o\sigma} \overline{“A”} \supset Theorem_{o\sigma} \overline{“B”}$.

7109 Corollary. For any wff \mathbf{A} , $\vdash Theorem_{o\sigma} \overline{“A”} \wedge Theorem_{o\sigma} \overline{“\sim A”} \supset \sim Consis$.

7110 Gödel's Second Incompleteness Theorem. Let \mathcal{A} be any recursively axiomatized pure extension of \mathcal{Q}_0^∞ . If \mathcal{A} is consistent, then not $\vdash_{\mathcal{A}} Consis$.

7111 Lemma.

$$\begin{aligned} \vdash \dot{\forall} x_\sigma \dot{\forall} y_\sigma \blacksquare Theorem_{o\sigma} [\overline{“[[\mathcal{D}_{ooo}] * x_\sigma * \overline{2048} * y_\sigma * \overline{2048}]”}] \\ \supset \blacksquare Theorem_{o\sigma} x_\sigma \supset Theorem_{o\sigma} y_\sigma. \end{aligned}$$

7112 Corollary. For any wfss_o \mathbf{A} and \mathbf{B} , $\vdash \text{Theorem}_{oo} \overline{[\mathbf{A} \supset \mathbf{B}]} \supset \overline{\text{Theorem}_{oo} \overline{\mathbf{A}}} \supset \overline{\text{Theorem}_{oo} \overline{\mathbf{B}}}$.

7113 Corollary. If $\vdash \mathbf{A} \supset \mathbf{B}$, then $\vdash \text{Theorem}_{oo} \overline{\mathbf{A}} \supset \text{Theorem}_{oo} \overline{\mathbf{B}}$.

7114 Löb's Theorem. For any sentence \mathbf{D} , $\vdash \mathbf{D}$ if and only if $\vdash \text{Theorem}_{oo} \overline{\overline{\mathbf{D}}} \supset \mathbf{D}$.

7115 Corollary. If \mathbf{D} is any sentence such that $\vdash \mathbf{D} = \text{Theorem} \overline{\overline{\mathbf{D}}}$, then $\vdash \mathbf{D}$.

7200 Gödel-Rosser Incompleteness Theorem. [Rosser, 1936] Let \mathcal{A} be a consistent, recursively axiomatized pure extension of \mathcal{Q}_0^∞ . Let ∇_{oo} be

$$\begin{aligned} & [\lambda x_\sigma \dot{\forall} y_\sigma \bullet \text{Proof}_{oo\sigma}^{\mathcal{A}}[\overline{128} * x_\sigma * [\text{Num}_{\sigma\sigma} x_\sigma] * \overline{2048}]y_\sigma \supset \\ & \dot{\exists} z_\sigma \bullet z_\sigma \leq y_\sigma \wedge \text{Proof}_{oo\sigma}^{\mathcal{A}}[\overline{[\sim_{oo}]} * x_\sigma * [\text{Num}_{\sigma\sigma} x_\sigma] * \overline{[?]}]z_\sigma. \end{aligned}$$

Let $m = \nabla_{oo}$ and Δ_o be $[\nabla_{oo} \overline{m}]$. Neither $\vdash_{\mathcal{A}} \Delta_o$ nor $\vdash_{\mathcal{A}} \sim \Delta_o$.

7201 Corollary.

- (a) If \mathcal{A} is any consistent and complete extension of \mathcal{Q}_0^∞ , then \mathcal{A} is not recursively axiomatizable.
- (b) If \mathcal{M} is any model of \mathcal{Q}_0^∞ , the set of sentences true in \mathcal{M} is not recursively axiomatizable.
- (c) If \mathcal{M} is any model of \mathcal{Q}_0^∞ , the set of sentences true in \mathcal{M} is not recursive.

7202 Proposition. $\vdash_{\mathcal{Q}_0} \text{Inf}^{III} \supset \text{Axiom } 6$.

7203 Proposition. There is a standard model \mathcal{N} for $\{\text{Inf}^{III}\}$ such that \mathcal{N} is isomorphic to every standard model for $\{\text{Inf}^{III}\}$.

7204 Corollary. \mathcal{Q}_0^∞ is consistent.

7205. Theorem. \mathcal{Q}_0^∞ is essentially incomplete.

7206 Theorem. There is no recursively axiomatized pure extension \mathcal{B} of \mathcal{Q}_0 such that the theorems of \mathcal{B} are precisely the wfss_o which are valid in the standard sense.

7207 Corollary. There is no decision procedure for determining of an arbitrary whether it is valid in the standard sense.

7300 Theorem. If \mathcal{A} is any consistent pure extension of \mathcal{Q}_0^∞ , then \mathcal{A} is undecidable.

7301 Theorem. If \mathcal{B} is an extension of \mathcal{Q}_0 such that $\{\text{Axiom 6}\}$ is consistent with \mathcal{B} , then \mathcal{B} is undecidable.

7302 Corollary. \mathcal{Q}_0 is undecidable.

7303 Proposition. Let \mathcal{A} be an extension of \mathcal{Q}_0 (or of \mathcal{F}) which is recursively axiomatized and complete. Then \mathcal{A} is decidable.

7304 Proposition. If \mathcal{M} is any model for \mathcal{Q}_0^∞ , then every numerical relation which is representable in \mathcal{Q}_0^∞ is definable in \mathcal{M} .

7305 Tarski's Theorem. For any model \mathcal{M} of \mathcal{Q}_0^∞ ,
 $\{\text{"A" } | \text{ A is a sentence and } \mathcal{M} \models \text{A}\}$ is not definable in \mathcal{M} .

7306 Theorem. Let \mathcal{A} be a consistent pure extension of \mathcal{Q}_0^∞ . There is no wff $\mathbf{B}_{o\sigma}$ of \mathcal{A} such that for each sentence \mathbf{A} of \mathcal{A} , $\vdash_{\mathcal{A}} \mathbf{A} = \mathbf{B}_{o\sigma} \overline{\text{"A"}}$.

Bibliography

- [Anderson and Bledsoe, 1970] Robert Anderson and W. W. Bledsoe. A Linear Format for Resolution with Merging and a New Technique for Establishing Completeness. *Journal of the ACM*, 17:525–534, 1970.
- [Andrews and Bishop, 1996] Peter B. Andrews and Matthew Bishop. On Sets, Types, Fixed Points, and Checkerboards. In Pierangelo Miglioli, Ugo Moscato, Daniele Mundici, and Mario Ornaghi, editors, *Theorem Proving with Analytic Tableaux and Related Methods. 5th International Workshop. (TABLEAUX '96)*, volume 1071 of *Lecture Notes in Artificial Intelligence*, pages 1–15, Terrasini, Italy, May 1996. Springer-Verlag.
- [Andrews *et al.*, 1984] Peter B. Andrews, Dale A. Miller, Eve Longini Cohen, and Frank Pfenning. Automating Higher-Order Logic. In W. W. Bledsoe and D. W. Loveland, editors, *Automated Theorem Proving: After 25 Years*, Contemporary Mathematics series, vol. 29, pages 169–192. American Mathematical Society, 1984.
- [Andrews *et al.*, 1996] Peter B. Andrews, Matthew Bishop, Sunil Issar, Dan Nesmith, Frank Pfenning, and Hongwei Xi. TPS: A Theorem Proving System for Classical Type Theory. *Journal of Automated Reasoning*, 16:321–353, 1996.
- [Andrews, 1963] Peter B. Andrews. A Reduction of the Axioms for the Theory of Propositional Types. *Fundamenta Mathematicae*, 52:345–350, 1963.
- [Andrews, 1965] Peter B. Andrews. *A Transfinite Type Theory with Type Variables*. Studies in Logic and the Foundations of Mathematics. North-Holland, 1965.
- [Andrews, 1971] Peter B. Andrews. Resolution in Type Theory. *Journal of Symbolic Logic*, 36:414–432, 1971.

- [Andrews, 1972a] Peter B. Andrews. General Models and Extensionality. *Journal of Symbolic Logic*, 37:395–397, 1972.
- [Andrews, 1972b] Peter B. Andrews. General Models, Descriptions, and Choice in Type Theory. *Journal of Symbolic Logic*, 37:385–394, 1972.
- [Andrews, 1974a] Peter B. Andrews. Provability in Elementary Type Theory. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 20:411–418, 1974.
- [Andrews, 1974b] Peter B. Andrews. Resolution and the Consistency of Analysis. *Notre Dame Journal of Formal Logic*, 15(1):73–84, 1974.
- [Andrews, 1981] Peter B. Andrews. Theorem Proving via General Matings. *Journal of the ACM*, 28:193–214, 1981.
- [Andrews, 1989] Peter B. Andrews. On Connections and Higher-Order Logic. *Journal of Automated Reasoning*, 5:257–291, 1989.
- [Andrews, 2001] Peter B. Andrews. Classical Type Theory. In Alan Robinson and Andrei Voronkov, editors, *Handbook of Automated Reasoning*, volume 2, chapter 15, pages 965–1007. Elsevier Science, 2001.
- [Appel and Haken, 1976] K. Appel and W. Haken. Every Planar Map is Four Colorable. *Bulletin of the American Mathematical Society*, 82:711–712, 1976.
- [Baader and Siekmann, 1994] Franz Baader and Jörg Siekmann. Unification theory. In Dov M. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, pages 41–125. Oxford University Press, 1994.
- [Baader and Snyder, 2000] Franz Baader and Wayne Snyder. Unification theory. In Alan Robinson and Andrei Voronkov, editors, *Handbook of Automated Reasoning*. Elsevier Science, 2000.
- [Bernays and Schönfinkel, 1928] P. Bernays and M. Schönfinkel. Zum Entscheidungsproblem der mathematischen Logik. *Mathematische Annalen*, 99:342–372, 1928.
- [Beth, 1953] Evert W. Beth. On Padoa’s Method in the Theory of Definition. *Indag. Math.*, 15:330–339, 1953.
- [Beth, 1959] E. W. Beth. *The Foundations of Mathematics*. North-Holland Publishing Co., 1959.

- [Bibel, 1987] Wolfgang Bibel. *Automated Theorem Proving*. Vieweg, Braunschweig, second edition, 1987.
- [Boolos, 1998] George Boolos. *Logic, logic, and logic*. Harvard University Press, 1998.
- [Buss, 1994] Samuel R. Buss. On Gödel's Theorem on Lengths of Proofs I: Number of Lines and Speedup for Arithmetic. *Journal of Symbolic Logic*, 59:737–756, 1994.
- [Church, 1936a] Alonzo Church. A Note on the Entscheidungsproblem. *Journal of Symbolic Logic*, 1:40–41, 1936. Correction *ibid.*, 101–102.
- [Church, 1936b] Alonzo Church. An Unsolvable Problem of Elementary Number Theory. *Amer. J. Math.*, 58:345–363, 1936. Reprinted in [Davis, 1965].
- [Church, 1940] Alonzo Church. A Formulation of the Simple Theory of Types. *Journal of Symbolic Logic*, 5:56–68, 1940.
- [Church, 1956] Alonzo Church. *Introduction to Mathematical Logic*. Princeton University Press, Princeton, N.J., 1956.
- [Craig, 1957] W. Craig. Linear Reasoning. A New Form of the Herbrand-Gentzen Theorem. *Journal of Symbolic Logic*, 22:250–268, 1957.
- [Davis *et al.*, 1976] M. Davis, Y. Matijasevic, and J. Robinson. Hilbert's Tenth Problem Diophantine Equations: Positive Aspects of a Negative Solution. In *Proc. Sympos. Pure Math*, volume 28, pages 323–378, 1976.
- [Davis, 1965] M. Davis. *The Undecidable*. Raven Press, Hewlett, N. Y., 1965.
- [Degtyarev and Voronkov, 1998] Anatoli Degtyarev and Andrei Voronkov. What You Always Wanted to Know About Rigid E-Unification. *Journal of Automated Reasoning*, 20:47–80, 1998.
- [Dreben and Denton, 1966] Burton Dreben and John Denton. A Supplement to Herbrand. *Journal of Symbolic Logic*, 31:393–398, 1966.
- [Dreben and Goldfarb, 1979] Burton Dreben and Warren D. Goldfarb. *The Decision Problem: Solvable Classes of Quantificational Formulas*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1979.

- [Dreben *et al.*, 1963] Burton Dreben, Peter Andrews, and Stål Aanderaa. False Lemmas in Herbrand. *Bulletin of the American Mathematical Society*, 69:699–706, 1963.
- [Feferman, 1964] Solomon Feferman. Systems of Predicative Analysis. *Journal of Symbolic Logic*, 29:1–30, 1964.
- [Frege, 1879] Gottlob Frege. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle, 1879. Translated in [van Heijenoort, 1967].
- [Gallier and Snyder, 1989] Jean H. Gallier and Wayne Snyder. Complete Sets of Transformations for General *E*-Unification. *Theoretical Computer Science*, 67:203–260, 1989.
- [Gallier *et al.*, 1992] Jean H. Gallier, Paliath Narendran, Stan Raatz, and Wayne Snyder. Theorem Proving Using Equational Matings and Rigid *E*-Unification. *Journal of the ACM*, 39:377–429, 1992.
- [Gandy, 1956] R. O. Gandy. On the Axiom of Extensionality - Part I. *Journal of Symbolic Logic*, 21:36–48, 1956.
- [Gentzen, 1935] G. Gentzen. Untersuchungen über das Logische Schließen I und II. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1935. Translated in [Gentzen, 1969].
- [Gentzen, 1969] G. Gentzen. Investigations into Logical Deductions. In M. E. Szabo, editor, *The Collected Papers of Gerhard Gentzen*, pages 68–131. North-Holland Publishing Co., Amsterdam, 1969.
- [Gleason, 1966] Andrew M. Gleason. *Fundamentals of Abstract Analysis*. Addison Wesley, 1966.
- [Gödel, 1930] Kurt Gödel. Die Vollständigkeit der Axiome des logischen Funktionenkalküls. *Monatsh. Math. Phys.*, 37:349–360, 1930.
- [Gödel, 1931] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatsh. Math. Phys.*, 38:173–198, 1931.
- [Gödel, 1936] Kurt Gödel. Über die Länge von Beweisen. *Ergebnisse eines Mathematischen Kolloquiums*, 7:23–24, 1936. Translated in [Gödel, 1986], pp. 396–399.

- [Gödel, 1986] Kurt Gödel. *Collected Works, Volume I*. Oxford University Press, 1986.
- [Goldson and Reeves, 1993] Doug Goldson and Steve Reeves. Using Programs to Teach Logic to Computer Scientists. *Notices of the American Mathematical Society*, 40:143–148, 1993.
- [Goldson *et al.*, 1993] Douglas Goldson, Steve Reeves, and Richard Bornat. A Review of Several Programs for the Teaching of Logic. *The Computer Journal*, 36:373–386, 1993.
- [Gordon, 1986] Mike Gordon. Why higher-order logic is a good formalism for specifying and verifying hardware. In G. J. Milne and P. A. Subrahmanyam, editors, *Formal Aspects of VLSI Design*, pages 153–177. North-Holland, 1986.
- [Hanna and Daeche, 1985] F. K. Hanna and N. Daeche. Specification and Verification using Higher-Order Logic. In Koomen and Moto-oka, editors, *Computer Hardware Description Languages and their Applications*, pages 418–433. North Holland, 1985.
- [Hanna and Daeche, 1986] F. K. Hanna and N. Daeche. Specification and Verification Using Higher-Order Logic: A Case Study. In G. J. Milne and P. A. Subrahmanyam, editors, *Formal Aspects of VLSI Design*, pages 179–213. North-Holland, 1986.
- [Hatcher, 1968] William S. Hatcher. *Foundations of Mathematics*. W. B. Saunders Co., 1968.
- [Hazen, 1983] A.P. Hazen. Ramified Type Theories. In D.M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume I. Reidel, Dordrecht, 1983.
- [Henkin, 1949] Leon Henkin. The Completeness of the First-Order Functional Calculus. *Journal of Symbolic Logic*, 14:159–166, 1949.
- [Henkin, 1950] Leon Henkin. Completeness in the Theory of Types. *Journal of Symbolic Logic*, 15:81–91, 1950.
- [Henkin, 1952] Leon Henkin. A Problem Concerning Provability. *Journal of Symbolic Logic*, 17:160, 1952.
- [Henkin, 1960] Leon Henkin. On Mathematical Induction. *American Mathematical Monthly*, 67:323–338, 1960.

- [Henkin, 1963a] Leon Henkin. An Extension of the Craig-Lyndon Interpolation Theorem. *Journal of Symbolic Logic*, 28:201–216, 1963.
- [Henkin, 1963b] Leon Henkin. A Theory of Propositional Types. *Fundamenta Mathematicae*, 52:323–344, 1963.
- [Herbrand, 1930] Jacques Herbrand. Recherches sur la théorie de la démonstration. *Travaux de la Société des Sciences et des Lettres de Varsovie, Classe III Sciences Mathématiques et Physiques*, 33, 1930. Translated in [Herbrand, 1971].
- [Herbrand, 1971] Jacques Herbrand. *Logical Writings*. Harvard University Press, 1971. Edited by Warren D. Goldfarb.
- [Hindley, 1997] J. Roger Hindley. *Basic Simple Type Theory*. Cambridge University Press, 1997.
- [Hintikka, 1955] K. J. J. Hintikka. Notes on Quantification Theory. *Soc. Sci. Fenn. Comment. Phys. Math*, 17, 1955.
- [Huet and Oppen, 1980] Gérard Huet and D. C. Oppen. Equations and Rewrite Rules: a Survey. In Ronald V. Book, editor, *Formal Language Theory: Perspectives and Open Problems*, pages 349–405. Academic Press, 1980.
- [Huet, 1975] Gérard P. Huet. A Unification Algorithm for Typed λ -Calculus. *Theoretical Computer Science*, 1:27–57, 1975.
- [Jacobs, 1999] Bart Jacobs. *Categorical Logic and Type Theory*. Elsevier, 1999.
- [Jensen and Pietrzykowski, 1976] D.C. Jensen and T. Pietrzykowski. Mechanizing ω -Order Type Theory Through Unification. *Theoretical Computer Science*, 3:123–171, 1976.
- [Jouannaud and Kirchner, 1991] J.-P. Jouannaud and C. Kirchner. Solving Equations in Abstract Algebras: A Rule-Based Survey of Unification. In J.-L. Lassez and G. Plotkin, editors, *Computational logic : essays in honor of Alan Robinson*. MIT Press, Cambridge, MA, 1991.
- [Kemeny, 1949] John G. Kemeny. *Type Theory vs. Set Theory*. PhD thesis, Princeton University, 1949. (abstract in [Kemeny, 1950]).
- [Kemeny, 1950] John G. Kemeny. Type theory vs. set theory. *Journal of Symbolic Logic*, 15:78, 1950. (abstract).

- [Kleene, 1952] S. C. Kleene. *Introduction to Metamathematics*. Van Nostrand, 1952.
- [Leivant, 1994] Daniel Leivant. Higher Order Logic. In Dov M. Gabbay, C.J. Hogger, and J.A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 2, pages 229–321. Oxford University Press, 1994.
- [Löb, 1955] M.H. Löb. Solution to a Problem of Leon Henkin. *Journal of Symbolic Logic*, 20:115–118, 1955.
- [Löwenheim, 1915] Leopold Löwenheim. Über Möglichkeiten im Relativkalkül. *Mathematische Annalen*, 76:xque447–470, 1915.
- [Lyndon, 1959] Roger C. Lyndon. An Interpolation Theorem in the Predicate Calculus. *Pacific J. Math*, 9:129–142, 1959.
- [Marshall and Chuaqui, 1991] M. Victoria Marshall and Rolando Chuaqui. Sentences of type theory: the only sentences preserved under isomorphisms. *Journal of Symbolic Logic*, 56:932–948, 1991.
- [Mendelson, 1964] E. Mendelson. *Introduction to Mathematical Logic*. D. Van Nostrand Company Ltd., Princeton, N.J., 1964.
- [Mendelson, 1987] Elliott Mendelson. *Introduction to Mathematical Logic*. Wadsworth & Brooks/Cole Advanced Books & Software, third edition, 1987.
- [Meredith, 1953] Carew A. Meredith. Single Axioms for the Systems (C, N) , (C, O) , and (A, N) of the Two-Valued Propositional Calculus. *J. Comput. Systems*, 1:155–164, 1953.
- [Miller, 1987] Dale A. Miller. A Compact Representation of Proofs. *Studia Logica*, 46(4):347–370, 1987.
- [Moore, 1982] Gregory H. Moore. *Zermelo's Axiom of Choice : Its Origins, Development, and Influence*. Springer-Verlag, 1982.
- [Paterson and Wegman, 1978] M.S. Paterson and M.N. Wegman. Linear Unification. *Journal of Computer and System Sciences*, 16:158–167, 1978.
- [Prawitz, 1965] Dag Prawitz. *Natural Deduction*. Almqvist & Wiksell, 1965.
- [Prawitz, 1968] Dag Prawitz. Hauptsatz for Higher Order Logic. *Journal of Symbolic Logic*, 33:452–457, 1968.

- [Quine, 1950] W. V. Quine. On Natural Deduction. *Journal of Symbolic Logic*, 15:93–102, 1950.
- [Quine, 1963] Willard Van Orman Quine. *Set Theory and Its Logic*. Belknap Press of Harvard University Press, 1963.
- [Rasiowa, 1949] H. Rasiowa. Sur un certain système d’axiomes du calcul des propositions. *Norsk matematisk Tidsskrift*, 31:1–3, 1949.
- [Robinson, 1965] J. A. Robinson. A Machine-Oriented Logic Based on the Resolution Principle. *Journal of the ACM*, 12:23–41, 1965.
- [Rosser, 1936] J. B. Rosser. Extensions of some Theorems of Gödel and Church. *Journal of Symbolic Logic*, 1:87–91, 1936.
- [Rubin and Rubin, 1985] Herman Rubin and Jean E. Rubin. *Equivalents of the Axiom of Choice, II*. North-Holland, 1985.
- [Russell, 1908] Bertrand Russell. Mathematical Logic as Based on the Theory of Types. *American Journal of Mathematics*, 30:222–262, 1908. Reprinted in [van Heijenoort, 1967], pp. 150–182.
- [Schönfinkel, 1924] M. Schönfinkel. Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, 92:305–316, 1924.
- [Schütte, 1950] K. Schütte. Schlussweisen Kalküle der Prädikatenlogik. *Mathematische Annalen*, 122:47–65, 1950.
- [Schütte, 1956] Kurt Schütte. Ein System des Verknüpfenden Schliessens. *Archiv für Mathematische Logik und Grundlagenforschung*, 2, 1956.
- [Shapiro, 1991] Stewart Shapiro. *Foundations without Foundationalism : A Case for Second-order Logic*. Clarendon Press, Oxford, 1991.
- [Sieg, 1997] Wilfried Sieg. Step by recursive step: Church’s analysis of effective calculability. *Bulletin of Symbolic Logic*, 3:154–180, 1997.
- [Siekmann, 1984] Jörg Siekmann. Universal Unification. In R. E. Shostak, editor, *Proceedings of the 7th International Conference on Automated Deduction*, volume 170 of *Lecture Notes in Computer Science*, pages 1–42, Napa, California, USA, 1984. Springer-Verlag.
- [Siekmann, 1989] Jörg Siekmann. Unification Theory. *Journal of Symbolic Computation*, 7:207–274, 1989.

- [Skolem, 1920] Thoralf Skolem. Logisch-kombinatorische Untersuchungen über die erfüllbarkeit oder beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen. *Skr. av Videnskapsselskapet i Kristiana, I. Matem.-natur klasse*, 4, 1920.
- [Skolem, 1928] Thoralf Skolem. Über die Mathematische Logik. *Norse matematisk tidsskrift*, 10:125–142, 1928. Translated in [van Heijenoort, 1967, pp. 508–524].
- [Smullyan, 1963] Raymond M. Smullyan. A Unifying Principle in Quantification Theory. *Proceedings of the National Academy of Sciences, U.S.A.*, 49:828–832, 1963.
- [Smullyan, 1968] R. M. Smullyan. *First-Order Logic*. Springer-Verlag, Berlin, 1968.
- [Snyder, 1991] Wayne Snyder. *A Proof Theory for General Unification*. Birkhäuser Boston, 1991.
- [Statman, 1978] Richard Statman. Bounds for Proof Search and Speed-up in the Predicate Calculus. *Annals of Mathematical Logic*, 15:225–287, 1978.
- [Takahashi, 1967] Moto-o Takahashi. A Proof of Cut-Elimination Theorem in Simple Type Theory. *Journal of the Mathematical Society of Japan*, 19:399–410, 1967.
- [Takeuti, 1953] Gaisi Takeuti. On a Generalized Logic Calculus. *Japanese Journal of Mathematics*, 23:39–96, 1953. Errata: *ibid*, vol. 24 (1954), 149–156.
- [Takeuti, 1958] Gaisi Takeuti. Remark on the fundamental conjecture of GLC. *Journal of the Mathematical Society of Japan*, 10:44–45, 1958.
- [Tarski, 1936] Alfred Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1:261–405, 1936. translated in [Tarski, 1956].
- [Tarski, 1956] A. Tarski. The Concept of Truth in Formalized Languages. In J. H. Woodger, editor, *Logic, Semantics, Metamathematics*, pages 152–278. Oxford University Press, Oxford, 1956.
- [Turing, 1948] A. M. Turing. Practical Forms of Type Theory. *Journal of Symbolic Logic*, 13:80–94, 1948.

- [van Benthem and Doets, 1983] J.F.A.K. van Benthem and K. Doets. Higher-Order Logic. In D. M. Gabbay and F. Günthner, editors, *Handbook of Philosophical Logic*, volume I, pages 275–329. Reidel, Dordrecht, 1983.
- [van Heijenoort, 1967] Jean van Heijenoort. *From Frege to Gödel. A Source Book in Mathematical Logic 1879–1931*. Harvard University Press, Cambridge, Massachusetts, 1967.
- [van Vaalen, 1975] Jophien van Vaalen. An Extension of Unification to Substitutions with an Application to Automatic Theorem Proving. In *Proceedings of the Fourth International Joint Conference on Artificial Intelligence*, pages 77–82, Tbilisi, Georgia, 1975.
- [Whitehead and Russell, 1913] Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*. Cambridge University Press, Cambridge, England, 1913. 3 volumes; first edition 1913, second edition 1927.
- [Wolfram, 1993] D. A. Wolfram. *The Clausal Theory of Types*, volume 21 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1993.

List of Figures

1.1	Truth table for \sim	7
1.2	Truth tables for \wedge , \vee , \supset , \equiv , $\not\equiv$	8
1.3	Truth tables for $[(p \supset q) \wedge (q \supset p)]$ and $[p \equiv q]$	9
1.4	Truth table for $\not\equiv$	10
1.5	$[\sim p \vee q]$ and $[p \supset q]$	14
1.6	$\sim [\sim p \vee \sim q]$ and $[p \wedge q]$	14
1.7	Plan for Chapter 1	17
1.8	A simple tautology	34
1.9	A short method of verifying a tautology	35
1.10		
	Table for Axiom 1	43
1.11		
	Table for Axiom 2	43
1.12		
	Table for Axiom 3	43
1.13	Binary truth functions.	45
1.14	\vee expressed in terms of \supset	47
1.15	Conjunctions of literals for dnf of $\sim \blacksquare [q \equiv r] \supset p$	49
1.16	dnf of $p \equiv q$	52
1.17	cnf of $p \equiv q$	52
1.18	Vpform for W1401: $[(p \vee q) \wedge \sim r \wedge \blacksquare \sim q \vee s] \vee \blacksquare \sim p \wedge \blacksquare r \vee \sim s$	52
1.19	Vpform for \sim W1402.	58
1.20	Vpform for \sim W1403.	59
1.21	Proof of 1502	65
2.1	Statements about voting in symbolic form and in English.	75
2.2	General advice on methods of proof	105
3.1	A tree	161
3.2	A semantic tableau	163

3.3	An abbreviated tableau	164
3.4	Modus ponens in a tableau	164
3.5	Flow chart for the unification algorithm	186
4.1	The tablet	190
4.2	Oren's translation of the tablet	191
4.3	Nero's translation of the tablet	191
4.4	Dual Pairs	192
5.1	Isomorphism condition (3)	246
5.2	Assignments into \mathcal{N} and \mathcal{M}	246
6.1	$(\mu z \leq y)r(z)$	296
7.1	Diagram for proof of 7203	328

Index

- \square , 162
 \Leftrightarrow , 44
 \Rightarrow , 44
 $*$, 298, 299, 313
 \bar{n} , 266
 \cap , 231
 \cup , 231
 \cap , 231
 \circ , 118, 351
 \cup , 127, 231, 352
 \div , 293
 $\dot{+}$, 288
 \downarrow , 51, 192, 347
 \leq , 274
 $=$, 142, 212, 235, 293, 342
 $\overline{\overline{p_{\alpha\beta}}}$, 258
 \neq , 212
 \equiv , 9, 14, 44, 212
 \exists_1 , 146, 233, 342
 $\dot{\exists}$, 260
 \exists , 79, 80, 99, 229, 360
 \exists' , 111
 $\exists \text{ Gen}$, 97, 212, 229
 $\exists I$, 100
 $\exists \text{ Rule}$, 99
 $\exists \mathbf{x} \in \mathbf{P}$, 79
 $\exists G$, 153, 157
 $\dot{\forall}$, 260
 \forall , 76, 77, 79, 81, 208, 212, 342
 $\forall \mathbf{x} \in \mathbf{P}$, 79
 $\forall G$, 153, 157
 $\forall I$, 92, 153, 221, 349, 358
- \leq , 269–273, 293, 365
 $<$, 206, 293, 365
 $|$, 44, 51, 192, 347
 \models , 35, 117, 239
 \natural , 86
 $\not\equiv$, 10, 192
 $\not\subset$, 192
 $\not\models$, 192
 $+$, 284
 $\#$, 231
 \sim , 7, 11, 33, 38, 40, 44, 49, 76, 212
 \blacksquare , 15
 \star , 166
 $\star\star$, 314, 368
 \subset , 44, 192
 \subseteq , 231
 \supset , 8, 14, 44, 212, 227, 342, 359
 $\supset \forall$, 87, 93, 227, 359
 $(\varphi : \mathbf{x}/z)$, 238
 \vdash , 24, 69, 86, 91, 92, 214, 229, 259, 349
 \vee , 8, 11, 33, 44, 76, 212, 342
 $\dot{\vee}$, 68
 \wedge , 7, 14, 44, 77, 212
 $\vec{x_k}$, 289
 $\times_{\sigma\sigma\sigma}$, 284
 $\mathbf{F}_{\alpha\alpha}^n \mathbf{A}_\alpha$, 266
 $\{\mathbf{F}_\alpha \mid \mathbf{A}_o\}$, 231
 0_σ , 260
- abbreviation, 14
absolutely consistent, 35

- absolutely inconsistent, 32, 35
- abstract consistency class, 130
- AC, 236
- additive abelian group, 85
- agree off, 115
- agree on, 115
- algorithm, 287
- α , 219
- α -conversion, 219, 342
- Alphabetic Change of Bound Variables, 91, 153, 156, 219, 342, 349
- and, 7
- assertion, 152
- assignment, 115, 238
- association to the left, 15
- atomic wff, 77
- axiom, 23, 84, 213, 342
- Axiom 6, 258
- Axiom of Choice, 236, 245
- axiom of infinity, 258, 266
- axiom schema, 23
- axiomatic method, 21
- axiomatic set theory, 319
- basic tautologies, 59
- β , 219
- β -contraction, 219, 342
- β -conversion, 219
- β -expansion, 219, 342
- β -normal form, 232
- β -reduction, 219
- Beth's Definability Theorem, 199
- bound, 81, 213
- bracket, 14
- but, 7
- \mathcal{C} , 341
- c-instance, 177
- Cantor's Theorem, 254
- card, 128, 248
- cardinal number, 258
- cardinality, 128, 257, 258
- Cases, 153
- categorical, 256, 326
- characteristic function, 268
- Church's Theorem, 86
- Church's Thesis, 287, 302
- closed, 13, 67, 81, 162
- closed under subsets, 129
- cnf, 48
- compact, 67
- compactness, 64, 136, 138, 142, 253, 330
- Compactness Theorem, 64, 136, 179, 181
- complete, 51, 142, 248, 334
- completeness, 135
- Completeness Theorem, 37, 135, 158, 163, 173, 253
- composition, 285
- compound Herbrand instance, 179
- compound instance, 177
- concatenate, 11
- conjunct, 48
- conjunction, 7, 48
- Conjunction Elimination, 162, 173
- Conjunction Introduction, 156
- conjunctive normal form, 48
- connection method, 177
- consequence, 142
- conservative expansion, 127
- conservative extension, 127, 243
- Consistency Theorem, 36, 86, 243
- consistent, 35, 86, 121, 126, 136, 243, 329
- constant, 76
- contradictory, 117
- countable, 128, 248, 251
- countable model, 244

- countermodel, 117
Craig-Lyndon Interpolation Theorem, 196
Cut, 68, 160, 173
Cut-Elimination, 160
cwff, 81, 213
- decision problem, 38, 177
Ded, 93, 153, 349
Deduction Theorem, 93, 349
definable, 335
define implicitly, 199
DeMorgan's laws, 48
denotation, 238
denote, 73
derived rules of inference, 27
descriptions, 343
designate, 42
disagreement set, 185
disjunct, 48
disjunction, 8, 10, 48
Disjunction Elimination, 162
Disjunction Rules, 156, 172
disjunctive component, 156, 172
disjunctive normal form, 48
disjunctively valid, 67
dnf, 48
domain of individuals, 80, 114
Double-Negation Elimination, 162
dual, 40, 192
duality, 192
- effectively computable, 287
elementarily equivalent, 146
elementary theory, 85
elementary type theory, 343
elementary wff, 77
elusive, 338
equality, 142
equality-interpretation, 146
- equality-model, 146
equipollent, 257
equivalence, 9
equivalent, 9, 127
equivalent as clauses, 68
equivalent systems, 343
essentially existential, 166
essentially incomplete, 329
essentially universal, 166
ETPS, xiv
exclusive disjunction, 8, 10
existential closure, 118
existential constant, 166
Existential Generalization, 153, 157
Existential Instantiation, 162
Existential Rule, 99
expansion, 126, 248
expansion proof, 320
explicit definition, 199, 290
explicitly definable, 290
expression, 184
extension, 127
extensionality, 343
extensionally complete, 248
- F_o , 212
 \mathcal{F} , 73, 84
 $\mathcal{F}^=$, 142
false, 117
 $Finite_{o(o)}$, 260
finite, 260
finite character, 129
finite intersection property, 67
finite model, 244
finite number, 260
first-order theorem, 243
first-order theory, 85
formation rule, 11, 77
formation sequence, 19
formula, 11

- Four Color Theorem, 136
 fourth-order logic, 201
 frame, 238
 free, 81, 92, 152, 213
 free for, 82, 152, 213
 free on, 118
 frugal, 128, 251, 255
 full conjunctive normal form, 48
 full disjunctive normal form, 48

 Gödel's Completeness Theorem, 135, 243
 Gödel's Second Theorem, 322, 324, 369
 Gen, 84, 91
 general model, 238
 general recursive, 285
 Generalization, 84, 342
 generalize on, 84
 Gentzen's Hauptsatz, 160
 Gödel, 243, 320
 graph, 136
 Ground Resolution, 69
 ground resolvent, 68

 \mathcal{H} , 23
 η , 219
 η -contraction, 219
 η -conversion, 219
 η -expansion, 219
 η -reduction, 219
 hardware, x
 Henkin, 321
 Henkin's Completeness Theorem, 253, 362
 Henkin's Theorem, 251
 Herbrand expansion, 182
 Herbrand property, 182
 Herbrand universe, 179
 Herbrand's Theorem, 179, 356

 Hilbert, 302
 Hilbert's program, 302, 319
 Hilbert-style system, 151, 182
 horizontal path, 52
 Hyp, 91, 152, 162
 hypotheses, 24, 152

 $\iota_{(\iota(\alpha))}$, 211, 341
 ι , 202, 210
 identity relation, 238
 if, 8
 if and only if, 10
 iff, 10
 Implication, 8
 implies, 8
 improper symbol, 11
 inclusive disjunction, 8
 inconsistent with respect to negation, 35
 independent, 42, 123
 Indirect Proof, 104, 153, 351
 individual, 238
 individual variable, 75
 Induction on the Construction of a Wff, 12, 345
 Inf^I , 258
 Inf^{II} , 259
 Inf^{III} , 326
 infinitesimal, 138
 infix, 45
 initial function, 285
 initial object, 13
 interpolation sentence, 196
 interpretation, 114, 238
 IP, 104, 153, 351

 Kemeny, 320

 $\mathcal{L}(Q_0)$, 248
 $\mathcal{L}(\mathcal{F})$, 126
 Löb, 321

- Löb's Theorem, 322, 369
Löwenheim's Theorem, 135
Löwenheim-Skolem Theorem, 136
 λ , 47, 207, 211, 219, 238
 λ -conversion, 219
 λ -notation, 279
language, 126, 248
least-number operator, 279
length of proof, 243
less specified, 184
liar paradox, 313
linear ordering, 66
literal, 48, 125
logistic system, 24
- material implication, 8
mating method, 177
matrix, 111
meaning, 115
meta-language, 11
metatheorem, 11
method of horizontal paths, 57
method of vertical paths, 57, 58
minimization, 279
model, 117, 121, 136, 238
Modus Ponens, 23, 84, 86, 342
Modus Tollens, 86
monotonic, 25
more general, 184
MP, 23, 84, 91
- $\mathbb{N}_{o\sigma}$, 260
nand, 44
natural deduction, 152
natural number, 17
necessary, 8, 10
Neg., 153
negation, 7
Negation Elimination, 173
Negation Introduction, 156
- negation normal form, 51, 177
negation normal formula, 51
negation normal sentence, 177
negative, 196
negatively, 111
nnf, 51, 177
nns, 177
Non-equivalence, 10
nonmonotonic, 26
nonstandard, 254, 255
nonstandard analysis, 138
nonstandard model, 138
nonstandard number theory, 138
nor, 44
not, 7
nth-order arithmetic, 320
null formula, 152
numeral, 266
numerical function, 278
numerical relation, 279
- o*, 202, 210
object language, 11
only if, 8
operation, 13
or, 8
- $\Pi_{o(o\alpha)}$, 212
 \mathcal{P} , 5
parable, 312
paradox, 136, 313, 322
partial ordering, 66
partial recursive, 285
PCI, 18
Peano's Postulates, 258, 261, 266
Peirce's dagger, 44
PMI, 17
poset, 66
positive, 196
positively, 111

- postulate, 85
 predicate, 74
 prefix, 45, 111
 prenex normal form, 111
 preserve, 31
 primitive recursion, 281
 primitive recursive, 284
 primitive rule of inference, 27
 primitive symbol, 11, 76
 principal dual, 192
 Principle of Complete Induction,
 18
 Principle of Mathematical Induc-
 tion, 17
 projection, 285
 proof, 24, 152, 213
 proof from hypotheses, 91, 214
 proper expansion, 126
 proper substitution, 46, 81
 proper symbol, 11
 propositional calculus, 5, 6
 propositional connective, 6, 14
 propositional variable, 10, 11
 pure, 132, 134

 \mathcal{Q}_0^1 , 343
 \mathcal{Q}_0^∞ , 259
 \mathcal{Q}_0^n , 329
 $\mathbf{Q}_{((\alpha\alpha)\alpha)}$, 211, 342
 \mathcal{Q}_0 , 209, 210
 quantifier, 75

 $R_{\sigma\sigma\sigma(\sigma\sigma\sigma)}$, 282
 ramified type theory, 202
 rectified, 111
 recursion operator, 281
 recursive, 285
 recursive function, 285
 recursively enumerable, 312
 refutation, 69

 regular, 279
 relativized quantifier, 79
 represent, 267
 representable, 267
 resolvent, 68
 Rosser, 324, 369
 Rule C, 153
 Rule of Cases, 104
 Rule of Generalization, 92
 rule of inference, 23, 84
 Rule P, 86, 93, 153, 227, 349
 Rule Q, 227, 359
 Rule R, 213
 Rule R', 214
 rules of inference, 342
 Russell, 202

 \sum^1 , 233
 $S_{\sigma\sigma}$, 260
 satisfiable, 116
 satisfies, 34, 116
 scope, 14, 79
 second-order logic, 201
 semantic tableau, 161
 semantical value, 115
 semantics, 17
 sentence, 81, 213
 Simp, 156
 simple Herbrand instance, 179
 simple instance, 177
 simple type theory, 202
 Simplification, 156, 172
 Skolem normal form, 167
 Skolem's Paradox, 136, 254
 Skolemized form, 167
 Smullyan's Unifying Principle, 129,
 132
 software, x
 sound, 35
 soundness, 146

- Soundness Theorem, 35, 240
special, 42
specification, x
standard model, 239
Statman, 243
stronger, 320
Substitution, 172
substitution, 16, 184, 342
substitution instance of a tautology, 85
Substitution-Value Theorem, 46, 347
Substitutivity of Implication, 88
subsume, 71
successor, 260
sufficient, 8, 10
sufficiently pure, 132
symmetry, 192
syntactical value, 77
syntactical variable, 11, 12, 23
syntax, 16
syntax language, 11
- T_o , 212
 T , 343
t-f contradiction, 179
t-f satisfies, 179
tautologous, 226
tautology, 226
term, 73, 77
theorem, 24, 85, 213
theorem schema, 36
theory, 21, 85
third-order logic, 201
topological space, 67
total ordering, 66
Trans, 28, 86, 345
transfinite type theory, 202, 319
Transitive Law of Implication, 28, 86, 345
true, 117, 336
- Truth, 338
truth assignment, 178
truth value, 238
truth-functional contradiction, 179
truth-functionally satisfies, 179
type symbol, 201
- Unification Algorithm, 186
unifier, 184
unify, 183, 184
universal, 166
universal closure, 118
Universal Generalization, 153, 157
Universal Instantiation, 92, 153, 162, 173, 349
universe, 80, 114
unless, 10
unsatisfiable, 117
- \mathcal{V}^M , 238
vacuous, 111
valid, 35, 117
value, 32, 115
variable, 76
variable-free, 130
verification, x
vertical path, 52
vertical path form, 52
vicious-circle principle, 202
vpform, 52
- Weaken, 156
Weakening, 156
well-formed formula, 11
well-formed part, 14, 80
wf part, 14, 80
wff, 11, 12, 77
wff of propositional calculus, 44
Whitehead, 202
 ω -order logic, 202

zero, 260