# A NEW INTRODUCTION TO MODAL LOGIC

*G. E. Hughes*
*Late Professor of Philosophy*
*Victoria University of Wellington*

*M. J. Cresswell*
*Professor of Philosophy*
*Victoria University of Wellington*

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First Published 1996
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

Reprinted 1998, 2001, 2003, 2004

*Routledge is an imprint of the Taylor & Francis Group*



Typeset in Times by M.J. Cresswell

Printed and bound in Great Britain by
TJ International Ltd., Padstow, Cornwall



*British Library Cataloguing in Publication Data*
A catalogue record for this book is available from the British Library.

*Library of Congress Cataloguing in Publication Data*
A catalogue record for this book has been requested.

ISBN 0-415-12599-5 (hbk)
ISBN 0-415-12600-2 (pbk)

# CONTENTS

# PREFACE

Modal logic is the logic of necessity and possibility, of 'must be' and 'may be'. By this is meant that it considers not only truth and falsity applied to what is or is not so as things actually stand, but considers what would be so if things were different. If we think of how things are as the actual world then we may think of how things might have been as how things are in an alternative, non-actual but possible, state of affairs – or possible world. Logic is concerned with truth and falsity. In modal logic we are concerned with truth or falsity in other possible worlds as well as the real one. In this sense a proposition will be necessary in a world if it is true in all worlds which are possible relative to that world, and possible in a world if it is true in at least one world possible relative to that world. All this is explained in the first chapter of this book.

Our aim in this book is to introduce readers to modal logic, and we assume that to begin with the reader knows nothing of modal logic. We have attempted to make the book self contained so that it could even be tackled by someone who had not studied any logic at all. However, we anticipate that most readers will already know a little about the (non-modal) propositional and predicate calculi, and will be able to use this knowledge as a foundation for understanding modal logic.

This book is intended as a replacement for our earlier two books *An Introduction to Modal Logic* (Hughes and Cresswell, 1968, IML) and *A Companion to Modal Logic* (Hughes and Cresswell, 1984, CML) and we shall here say a little about the relation between it and the earlier books. Part I covers most of the ground covered in IML with two important changes. First, as in CML, we take the system K as basic rather than T. Second, as also in CML, we have (in Chapter 6) used the method of canonical models to prove completeness. We have retained (in Chapter 5) the method of modal conjunctive normal forms to prove the completeness of S5, but while (in Chapter 4) we have retained from IML the method of semantic diagrams for testing formulae, we have omitted the completeness proofs based on this method.

Part II covers a range of topics in modal propositional logic, most of which are also discussed in CML. In the present work we have attempted to be particularly sensitive to its role as an introduction. Thus, to take one example, our approach to finite models is one that we believe is easier to follow than the more standard method of filtrations which we used in CML. Although this part of the book may be seen as more of interest to specialists we have tried to present its topics in a way which should be easily accessible to the reader who has followed Part I. Part III of IML contained a survey of modal logic as it was in 1968. A comparable survey would be impossible today but we have attempted, in Chapter 11 of the present book, to give an outline of the more important developments in the earlier history of modal logic. Readers who need more may be referred to IML.

Part III was the most difficult to write. Modal predicate logic is rightly regarded as the most philosophically important branch of modal logic, and although this is a book on formal logic not philosophical logic, we have attempted to discuss topics which have a bearing on such important philosophical questions as what to say about things which exist in one world but not another or about things claimed to be identical but not necessarily so. Unfortunately, the semantics of modal predicate logic is extremely complicated, and while we have tried to make our discussions as approachable as we can, we are conscious of the burden imposed on the reader. All we can say is that we have attempted to set out all technical material so that with patience a reader should be able to follow every proof without requiring more than is in this book.

George Hughes died on 4 March 1994. At the time of his death we had completed the first five chapters. Chapter 6 and most of Part II has been adapted from CML, and we had discussed many issues in that area. In preparing the manuscript I have endeavoured, to the best of my ability, to write it as a joint work and present it in a style as close as I can to what would have emerged had George Hughes lived to see its completion. It is in Part III that I have felt the greatest lack of his collaboration, and I am grateful especially to Rob Goldblatt and Edwin Mares here in Wellington who have looked at and commented on many passages. We would also thank various readers, and colleagues from around the world, whose 'wish lists' we have not always been able to take as much note of as they would like.

Our department secretary, Debbie Luyinda, put the initial manuscript into the computer, so that we could then play with it, and we would thank her for this, at times frustrating, work.

Wellington, New Zealand
January 1995



# Part I

# BASIC MODAL PROPOSITIONAL LOGIC

# 1

# THE BASIC NOTIONS

In this chapter we introduce the basic notions of modal propositional logic. Modal logic is based upon the 'ordinary' (two-valued) Propositional Calculus, and when we use the expression 'Propositional Calculus' (or the abbreviation 'PC') *simpliciter*, it is to this non-modal system of logic that we shall be referring.[1] The present chapter begins by outlining, in a very summary fashion, those elements of PC which we shall take for granted in what follows, and at the same time explains some of the terminology which we shall use throughout the book.

## The language of PC

We take as *primitive* (or undefined) symbols of PC the following:

A set of *letters:* $p$, $q$, $r$, ... (with or without numerical subscripts). We suppose ourselves to have an unlimited number of these.

The following four symbols: $\sim$, $\vee$, (, ).

Any symbol in the above list, or any sequence of such symbols, we call an *expression*. An expression is either a formula – more exactly a well-formed formula (wff) – or else it is not. We are concerned only with expressions which are well-formed formulae (wff). The following *formation rules* of PC specify which expressions are to count as wff:

**FR1** A letter standing alone is a wff.
**FR2** If $\alpha$ is a wff, so is $\sim\alpha$.
**FR3** If $\alpha$ and $\beta$ are wff, so is $(\alpha \vee \beta)$.

In these rules the symbols $\alpha$ and $\beta$ are used to stand indifferently for any expressions. Thus the meaning of FR2 is: the result of prefixing $\sim$ to any wff is itself a wff. Symbols used as $\alpha$ and $\beta$ are used here are known as metalogical variables. They are not among the symbols of the system (PC in this case), but are used in talking about the system.

Examples of wff are: $p$, $\sim q$, $\sim\sim\sim q$, $(p \vee \sim q)$, $((p \vee r) \vee \sim(q \vee \sim(\sim r \vee p)))$. For convenience, however, we allow ourselves to omit the outermost brackets round any complete wff (though not any subordinate part thereof). No ambiguity in interpretation or unclarity about what is permitted by the rules will result from this notational simplification.

## Interpretation

We interpret the letters as variables whose values are *propositions*. We shall usually call them *propositional variables*. We assume that the reader is familiar with the notion of a proposition, and shall not enter into the philosophical issues which this notion raises. Rough synonyms of 'proposition' are 'statement' and 'assertion', where these words are used to refer to *what is stated or asserted*, not to the *act of stating or asserting*. Every proposition is either true or false, and no proposition is both true and false. (Hence if something is neither true nor false, or is capable of being both true and false, it is not to count as a proposition in the present context.) Truth and falsity are said to be the *truth-values* of propositions.

Now it is possible to form more complex propositions out of simpler ones. E.g., out of the proposition that Brutus killed Caesar we can form the proposition that it is not the case that Brutus killed Caesar. This is a proposition which is true if the original proposition is false, and false otherwise. In general, putting 'it is not the case that' in front of a sentence will result in a sentence which expresses a proposition which is true if the original sentence expresses one which is false, and a false proposition if it does not.

Similarly, from the proposition that Brutus killed Caesar and the proposition that Cassius killed Caesar we may form the proposition that either Brutus killed Caesar or Cassius killed Caesar. This proposition will be true iff (if and only if) at least one of the original propositions is true, and therefore false iff both of these are false.

'It is not the case that' and 'either … or …', when used in the way we have just described, may be said to be *proposition-forming operators* on propositions, because they make new propositions out of old ones. The propositions on which such an operator operates are called its arguments.





If an operator requires only a single argument, as 'it is not the case that' does, it is said to be *monadic*; if, like 'either … or …', it requires two, it is said to be *dyadic*.

Our explanation of these operators, 'it is not the case that' and 'either… or …', showed that the truth-value of a proposition formed by means of either of them depends in every case only on the truth-value of the operator's argument or arguments. In other words, whenever we are given the truth-value of the argument or arguments, we can deduce the truth-value of the complex proposition. An operator which has this property is said to be a *truth-functional* operator, and the propositions it forms are said to be *truth-functions* of its arguments. Not all proposition-forming operators are of this kind. For example, given merely the truth or falsity of the proposition that Brutus killed Caesar we cannot deduce the truth or falsity of the proposition that Napoleon believed that Brutus killed Caesar; and given merely that two propositions are both true we cannot deduce from this either the truth or the falsity of the proposition that the first follows logically from the second (though if we are given that one proposition is false and another true, we *can* deduce from this that it is false that the first follows logically from the second). Hence although 'Napoleon believed that' and 'follows logically from' are proposition-forming operators on propositions (monadic and dyadic respectively), they are not truth-functional operators.

We interpret $\sim$ and $\vee$ as 'it is not the case that' and 'either … or …' respectively, in the senses we have explained, and we usually read them simply as 'not' and 'or'. $\sim$ so interpreted is called the *negation sign*; $\sim p$ is said to be the *negation* of $p$. Using 1 and 0 for the truth-values truth and falsity respectively, we can express the meaning we attach to $\sim$ in the following *basic truth-table for negation*:

| | $\sim$ |
|---|---|
| 1 | 0 |
| 0 | 1 |

Here the left-hand column tabulates the possible truth-values of a given proposition, and the right-hand column sets down the corresponding truth-values of the negation of that proposition. When interpreted in the way we have described, $\vee$ is known as the *disjunction sign* and its arguments are called *disjuncts*; $p \vee q$ is said to be the *disjunction* of $p$ and $q$. The basic truth-table for disjunction is:

| $\vee$ | 1 | 0 |
|---|---|---|
| 1 | 1 | 1 |
| 0 | 1 | 0 |

The possible truth-values of the first disjunct are tabulated in the leftmost vertical column and those of the second in the topmost horizontal row. The truth-value of their disjunction is found by reading across and down.

These basic truth-tables bring out clearly the truth-functional nature of the operators. In fact, not merely $\sim$ and $\vee$, but all operators in PC, are truth-functional and for this reason PC is sometimes called the theory of truth-functions. We said earlier that we interpret $p$, $q$, $r$, ... as variables whose values are propositions; but in view of the fact that the only feature of the arguments of the operators which is relevant to the truth-value of the complex propositions they form is their truth-value, it is equally satisfactory from a formal point of view to regard the variables as having as their range of values, not the whole infinite set of propositions, but simply the two truth-values 1 and 0.

**Further operators**

A number of other operators can be defined in terms of the primitive ones. We introduce three new operators, $\wedge$, $\supset$ and $\equiv$, though it would be possible to have several others as well. The definitions are:

[Def $\wedge$] $(\alpha \wedge \beta) =_{df} \sim(\sim\alpha \vee \sim\beta)$
[Def $\supset$] $(\alpha \supset \beta) =_{df} (\sim\alpha \vee \beta)$
[Def $\equiv$] $(\alpha \equiv \beta) =_{df} ((\alpha \supset \beta) \wedge (\beta \supset \alpha))$

In these definitions $\alpha$ and $\beta$ represent any wff of PC and the symbol '$=_{df}$' is read as 'is defined as'. The meaning of the first definition is that whenever we have a wff of the form $\sim(\sim\text{—} \vee \sim\text{—})$, where the blanks are filled by any wff we please, we can replace this wff by an expression which consists of the wff which filled the first blank followed by a $\wedge$ followed by the wff which filled the second blank, the whole being enclosed in brackets. Analogous explanations apply to the two other definitions. Similarly, we can expand any expression of the form on the left into the corresponding expression of the form on the right.

Expressions which can be transformed, by applying definitions, into wff as specified by the original formation rules, are themselves to count





as wff. When a wff contains no symbols except primitive ones it is said to be written in *primitive notation.* The definitions enable us to write all wff in primitive notation if we wish to do so.

**Interpretation of $\wedge$, $\supset$ and $\equiv$**

The interpretation we have already given to $\sim$ and $\vee$ will determine the interpretation we give to the operators defined in terms of them. Thus, we can calculate the truth-values of $p \wedge q$ for all possible truth-values of $p$ and $q$ by calculating the appropriate truth-values of the wff of which it is an abbreviation, viz. $\sim(\sim p \vee \sim q)$, and the basic truth-tables for $\sim$ and $\vee$ enable us to do this. It turns out that $p \wedge q$ will be true when both $p$ and $q$ are true, but false in all other cases. The basic truth-table for $\wedge$ will therefore be:

| $\wedge$ | 1 | 0 |
|---|---|---|
| 1 | 1 | 0 |
| 0 | 0 | 0 |

When $\wedge$ is so interpreted, it is called the *conjunction sign*; it may be read as 'and'. A wff formed with $\wedge$ is known as a *conjunction*, and the arguments are called *conjuncts*.

Similar considerations give the following basic truth-table for $\supset$:

| $\supset$ | 1 | 0 |
|---|---|---|
| 1 | 1 | 0 |
| 0 | 1 | 1 |

I.e. a proposition formed with $\supset$ is false when the first argument is true and the second false, but true in all other cases. When so interpreted, $\supset$ is known as the *(material) implication sign.* It may be read as '(materially) implies' or as 'if [the first argument], then [the second argument]'. The first argument is known as the *antecedent*, the second as the *consequent*. The precise relation of material implication to the various uses of the word 'if' in English raises complex questions into which we shall not enter here. It may plausibly be claimed, however, that material implication represents the truth-functional component in the meaning of 'if' in at least a great many of its standard uses.

The basic truth-table for $\equiv$ works out as:

| $\equiv$ | 1 | 0 |
|---|---|---|
| 1 | 1 | 0 |
| 0 | 0 | 1 |

I.e. a proposition formed with $\equiv$ is true when both arguments have the same truth-value, false when they have different truth-values. When so interpreted, $\equiv$ is known as the *(material) equivalence sign*. It may be read as 'is (materially) equivalent to', or as 'if and only if'.

Clearly these new operators, like the primitive ones, are truth-functional.

(We could have chosen other operators than $\sim$ and $\vee$ as primitive. Some authors, for example, take $\sim$ and $\wedge$ as primitive and define $\vee$ in terms of these. But whatever primitives we use, provided that all the operators can consistently be given the basic truth-tables listed above, the system of PC so obtained will be exactly equivalent to the one we have set down here.)

**Validity**

If we regard the variables, $p$, $q$, $r$, ... as taking the whole range of propositions as their values, we can say that a wff of PC becomes a proposition when all its variables are replaced by propositions. A wff is said to be *valid* iff the result of *every* such replacement is a true proposition. (It is assumed that the replacement is carried out uniformly, i.e. that two or more occurrences of the same variable are always replaced by the same proposition.) If, however, we speak instead of the variables taking simply the two truth-values 1 and 0 as their values, we shall say that a wff is valid iff it always has the value 1, no matter what truth-values are (uniformly) assigned to its variables. We shall normally choose to speak in this second way; since all the operators in PC are truth-functional, exactly the same formulae will turn out to be valid in each case. Simple examples of valid wff are $p \vee \sim p$ and $(p \wedge q) \supset p$. (A valid wff of PC is often called a *tautology* or a *PC-tautology*.)

A wff is said to be *unsatisfiable* iff it always has the value 0, no matter what truth-values are (uniformly) assigned to its variables. A simple example of an unsatisfiable wff is $p \wedge \sim p$. Many wff, such as $p \supset q$, are of course neither valid nor unsatisfiable.

Later in this chapter we shall extend this definition of validity to cover the formulae of modal logic, and to make the extended definition more





easily comprehensible we shall express it in the form of a parlour game. As a preliminary to this let us now consider how we might devise a simple game based on the definition of PC-validity which we have just mentioned. The game could take this form. We give a player a sheet of paper on which we have previously written a number of letters of the alphabet (preferably taken from the series, $p$, $q$, $r$, ... etc.). We shall refer to the player and the sheet as a *setting* of the *PC game*, or more succinctly a *PC-setting*. PC-settings will differ only in the list of letters on the sheet of paper.

We then call out to the player wff of PC, to which the player is to respond by either raising his or her hand or keeping it down. But each call must be appropriately prepared for, in that before a wff $\alpha$ is called we must have previously called all the formulae which occur as parts of $\alpha$, beginning with the variables. E.g., if $(\sim p \vee p)$ is to be called we must first call $p$, and then $\sim p$ and only then may we call $(\sim p \vee p)$. The player's instructions are as follows:

1. If a single letter (variable) is called, raise your hand if that letter is on the sheet; keep it down if it is not.

2. If $\sim\alpha$ is called (where $\alpha$ is a wff) raise your hand if you kept it down when $\alpha$ was called; keep it down if you raised it when $\alpha$ was called. (Remember that if $\sim\alpha$ has been appropriately prepared for, $\alpha$ must have already been called.)

3. If $(\alpha \vee \beta)$ is called, raise your hand if you raised it for $\alpha$ or for $\beta$; keep it down if you kept it down for both $\alpha$ and $\beta$.

Using the definitions of $\supset$, $\wedge$ and $\equiv$ we can easily derive rules for responding to formulae containing these operators. Alternatively we can transform all formulae into primitive notation before we begin. It might be worth stating the rule for $\supset$ explicitly:

3a. If $(\alpha \supset \beta)$ is called, raise your hand if you kept it down for $\alpha$ or raised it for $\beta$; keep it down if you raised it for $\alpha$ and kept it down for $\beta$.

It is not difficult to see that in any PC-setting the rules require the player to respond unambiguously to any PC formula, provided that it is appropriately prepared for. If the player in a PC-setting raises his or her hand when a PC wff $\alpha$ is called, we shall say that $\alpha$ is *successful* in that setting. Many formulae will be successful in some settings but not in others (depending of course on which letters appear on the sheet for a given setting). But there will be some formulae which will be successful in every PC- setting (e.g. $p \vee \sim p$). These we call *PC-successful*.

To make explicit what must be becoming an obvious parallel, let us call the sheet of variables an assignment of truth-values with the idea that a variable has the value 1 if it is on the sheet and 0 otherwise. On this understanding, when the player's hand is raised when a wff $\alpha$ is called it will mean that $\alpha$ has the value 1, and when the player's hand is kept down when $\alpha$ is called it will mean that $\alpha$ has the value 0. The rules 1, 2 and 3 for responding to formulae when thus translated exactly reflect the basic truth-tables for $\sim$ and $\vee$. A formula will be successful in a PC-setting iff it has the value 1 for the corresponding assignment of truth-values to its variables. And a formula will be PC-successful iff it is has the value 1 for every PC-assignment. I.e., the PC-successful wff are precisely those which are PC-valid.

Since for any wff $\alpha$ containing n variables we need only consider sheets which contain a selection (possibly all or possibly none) of those n variables (for clearly the responses to variables not in $\alpha$ cannot affect the response to $\alpha$), we can set out all the relevantly different PC-settings on $2^n$ sheets. So we could check whether $\alpha$ is valid by preparing such a set of sheets and calling $\alpha$ (with the appropriate preparatory calls) for each of them. This procedure can be codified by what is called the truth-table method of testing for PC-validity.

### Testing for validity: (i) the truth-table method

In this method of testing a PC formula, $\alpha$, for validity, all possible PC value-assignments, i.e. all assignments of truth-values to the propositional variables in $\alpha$, are tabulated, and for each such value-assignment, the basic truth-tables for the operators are used to calculate the truth-value of $\alpha$ as 1 or 0. The result is a column of ls and/or 0s. This column is known as the *truth-table* of the wff. If and only if it consists entirely of 1s, the wff is valid.

An example should make the procedure clear. Let $\alpha$ be $((p \supset q) \wedge r) \supset ((\sim r \vee p) \supset q)$. Here we have three distinct variables and therefore eight PC value-assignments. The construction of the truth-table proceeds as follows:





| $p$ | $q$ | $r$ | $((p \supset$ | $q) \wedge$ | $r) \supset$ | $((\sim r$ | $\vee\ p)$ | $\supset q)$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 |
| 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 |
| 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| | | | (1) | (2) | (6) | (3) | (4) | (5) |

The complete list of value-assignments is set down to the left of the vertical line. The columns to the right are numbered in the order in which they are obtained. Thus column (1), for $p \supset q$, is obtained from the columns under $p$ and $q$ by the basic truth-table for $\supset$; column (2) is obtained from (1) and the column under $r$, by the basic truth-table for $\wedge$; ... until finally column (6), the truth-table for the whole wff, is obtained from (2) and (5). Since (6) consists entirely of 1s $\alpha$ is PC-valid.

### Testing for validity: (ii) the Reductio method

A formula can usually be tested more expeditiously by trying to find a falsifying value-assignment for it. The *Reductio* method enables us to find such a value-assignment if there is one.

We begin by supposing that there is such an assignment for which $\alpha$ has 0. We express this supposition by writing 0 under the main operator of $\alpha$. From this supposition certain consequences follow, by the basic truth-tables, about the values which must be assigned to certain well-formed parts of $\alpha$; e.g., if $\alpha$ is of the form $\beta \supset \gamma$, it can only have 0 if $\beta$ has 1 and $\gamma$ has 0. From these new values certain other consequences follow in the same way, and so on, until finally we either (i) reach a consistent value-assignment to all the variables in $\alpha$ (in which case $\alpha$ is invalid), or (ii) find that we cannot reach such a consistent value-assignment (in which case $\alpha$ is valid).

As an example, let $\alpha$ be the formula we used to illustrate the truth-table method, viz. $((p \supset q) \wedge r) \supset ((\sim r \vee p) \supset q)$. We set out the whole working immediately and then explain it.

| ((p | ⊃ | q) | ∧ | r) | ⊃ | ((~ | r | ∨ | p) | ⊃ | q) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 1 | <u>1</u> | 0 | 1 | <u>0</u> | 1 | 0 | 0 | 0 |
| 9 | 4 | 8 | 2 | 5 | 1 | 11 | 12 | 6 | 10 | 3 | 7 |

The numerals under the truth-values indicate the order of the steps. Step 1 is the initial assignment of 0 to $\alpha$. Since $\alpha$ is of the form $\beta \supset \gamma$ if $\alpha$ has 0, $\beta$ must have 1 (step 2) and $\gamma$ must have 0 (step 3). The 1s at steps 4 and 5 are required by the table for $\wedge$ since $\beta$ is a conjunction and must have the value 1. The remaining steps should now be clear. We finally reach the conclusion (indicated by underlining) that if we are to have $\alpha$ with 0 $r$ must have both the value 1 and the value 0. Hence $\alpha$ can never have 0, and is therefore valid.

Other cases are sometimes not so simple. Suppose that $\alpha$ is the converse of the previous formula, viz. $((\sim r \vee p) \supset q) \supset ((p \supset q) \wedge r)$. Steps 1, 2 and 3 can proceed as before, but the values at steps 2 and 3 do not determine further values uniquely. We can however list exhaustively the alternatives left open at step 2 by the assumption that $((\sim r \vee p) \supset q)$ has 1, as follows:

| | ((~r | ∨ | p) | ⊃ | q) | ⊃ | ((p ⊃ q) | ∧ | r) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | 1 | | 0 | | 0 | |
| | | | | 2 | | 1 | | 3 | |
| (a) | | 1 | | 1 | 1 | 0 | | 0 | |
| (b) | | 0 | | 1 | 1 | 0 | | 0 | |
| (c) | | 0 | | 1 | 0 | 0 | | 0 | |

(a), (b) and (c) represent all the value-assignments to $(\sim r \vee p)$ and $q$ which are compatible with the truth of $((\sim r \vee p) \supset q)$. If each of these leads us to an inconsistency, $\alpha$ is valid; if even one of them is compatible with a consistent assignment to the variables, $\alpha$ is not valid. In fact (b) and (c) both lead to inconsistencies; but (a) does not – it is compatible with $q = 1$, $r = 0$ and $p = 1$ or 0. Hence the whole formula is not valid.

Provided we consider in this way all alternative value-assignments as the need arises, we can test the validity of any wff of PC whatever by the Reductio method. We shall make considerable use of this method in Chapter 4.

Each of the two methods we have described gives us an effective (i.e.





mechanical and finite) procedure for deciding of any given wff of PC whether it is valid or not. Another way of expressing this is by saying that each method gives us a *decision procedure* for PC.

**Some valid wff of PC**

We list here some valid PC wff which we shall use in the next few chapters. In some cases we give, in addition to a reference number, a name by which the formula is commonly known and an abbreviation by which we shall usually refer to it in this book.

PC1 $(p \wedge q) \supset p$
PC2 $(p \wedge q) \supset q$
PC3 $(p \supset q) \supset ((p \supset r) \supset (p \supset (q \wedge r)))$ [Law of Composition–Comp]
PC4 $p \supset (q \supset (p \wedge q))$ [Law of Adjunction–Adj]
PC5 $(p \supset q) \supset ((q \supset p) \supset (p \equiv q))$
PC6 $(p \supset q) \supset ((q \supset r) \supset (p \supset r))$ [Law of Syllogism–Syll]
PC7 $(p \supset (q \supset r)) \supset ((p \wedge q) \supset r)$ [Law of Importation–Imp]
PC8 $(p \supset q) \supset ((q \supset (r \supset s)) \supset ((p \wedge r) \supset s))$
PC9 $p \supset (p \vee q)$
PC10 $q \supset (p \vee q)$
PC11 $(p \supset q) \supset ((r \supset q) \supset ((p \vee r) \supset q))$
PC12 $p \equiv \sim\sim p$ [Law of Double Negation–DN]
PC13 $(p \vee q) \equiv \sim(\sim p \wedge \sim q)$
PC14 $(p \wedge q) \equiv \sim(\sim p \vee \sim q)$ [De Morgan Laws—DeM]
PC15 $(p \supset q) \equiv (\sim q \supset \sim p)$ [Law of Transposition–Transp]
PC16 $(p \vee q) \equiv (q \vee p)$
PC17 $(p \wedge q) \equiv (q \wedge p)$ [Commutative Laws—Comm]
PC18 $((p \vee q) \vee r) \equiv (p \vee (q \vee r))$
PC19 $((p \wedge q) \wedge r) \equiv (p \wedge (q \wedge r))$ [Associative Laws—Assoc]
PC20 $p \equiv (p \vee p)$
PC21 $p \equiv (p \wedge p)$

**Basic modal notions**

On p. 5 we called attention to the distinction between truth-functional and non-truth-functional operators, and we noted that all the operators which we use in PC are interpreted purely truth-functionally. In modal logic, however, we are going to be concerned in addition with a number of non-truth-functional concepts, and to express these we shall extend the language of PC by adding to it some new operators which we shall interpret in a non-truth-functional way.

To begin with, we shall add to the language of PC a new monadic operator, $L$, with the formation rule that if $\alpha$ is a wff, so is $L\alpha$. We shall call $L$ the *necessity operator*, and our intended interpretation of it is that it is to express, in the form of a proposition-forming operator on propositions, the notion which is commonly expressed by English words or phrases such as 'necessarily', 'must be', 'is bound to be'. In ordinary English such expressions, like the truth-functional 'not', are frequently found in the middle of a sentence rather than at the beginning; but just as it is possible, at the cost of a little artificiality, to replace an embedded 'not' by the phrase 'it is not the case that' at the beginning of the sentence, and thereby bring out more clearly its nature as an operator on propositions, so we can, for example, re-cast a sentence of the form 'A is bound to be B' as 'It is bound to be the case that A is B'. Necessity is called a *modal* notion, presumably because being necessarily true has been thought of as a *mode* or manner in which a proposition can be true.

We shall usually read $Lp$ as 'Necessarily $p$'. But in doing so we do not intend to claim that our use of $L$ will reflect all the standard English uses of 'necessarily' and the other expressions we have mentioned, any more than we could claim that the basic truth-table for conjunction provides an adequate analysis of all standard English uses of 'and'. On the other hand, we do not want to restrict its meaning to a single narrowly conceived sense of 'necessarily', etc. Very often, for example, when we say that something *must* be so, we can be taken to be claiming that it *is* so; and if we take $L$ to express 'must be' in this sense, we shall want to have it as a principle that whenever $Lp$ is true, so is $p$ itself. On the other hand there are uses of words such as 'must' and 'necessary' in which they express not what necessarily *is so* but rather what *morally ought to be so*; and if we interpret $L$ in accordance with these uses we shall want to allow the possibility that $Lp$ may be true but $p$ itself false, since people do not always do what they ought to do. As we shall see in the next chapter, it will prove possible to devise systems of modal logic which contain 'If $Lp$ then $p$' as a principle, and other systems which do not. In fact, one of the important features of modal logic is that out of the same basic material we can construct a variety of systems which reflect a variety of interpretations of $L$, within the range which can be indicated, somewhat loosely, by calling it a necessity operator. We shall even sometimes extend the interpretation of $L$ a little beyond these limits; for fruitful systems of logic have been inspired by the idea of taking the necessity





operator to mean, for example, 'It will always be the case that ...', 'It is known that ...' or 'It is provable that ...'. All this should become clearer as we proceed.

One thing that should be clear already, however, is that in any of the interpretations we have referred to, the necessity operator is not a truth-functional one: that is, the truth-value of $p$ itself is not always sufficient to determine the truth-value of $Lp$. Hence we cannot define $L$ in terms of any combination of the PC operators, and we therefore introduce it as a new primitive symbol.

Another notion which leads, in a parallel way, to a monadic non-truth-functional operator is one expressed by terms such as 'possibly', 'can be', 'may be'. We shall use $M$ as an operator with this meaning, and we shall usually read $Mp$ as 'possibly $p$'. If we already have $L$ in our logical language, however, we do not need to have $M$ as a new *primitive* symbol; for to say that it is possible that $p$ is equivalent to saying that it is not necessary that not-$p$, and we can therefore define $M\alpha$, for any $\alpha$, as $\sim L \sim \alpha$. Thus for every interpretation of $L$ there will be a corresponding interpretation of $M$: if $Lp$ means that $p$ is necessarily true, $Mp$ will mean that $p$ is possibly true, if $Lp$ means that it is morally obligatory that $p$, $Mp$ will mean that it is morally permissible that $p$ (not obligatory that not-$p$), if $Lp$ means that it will always be the case that $p$, $Mp$ will mean that it will sometime be the case that $p$, and so forth. (If we had chosen to take $M$ as primitive we could have defined $L$ as $\sim M \sim$. Whether to take $L$ or $M$ as primitive is a matter of taste. We shall continue to take $L$ as primitive and $M$ as defined.) Impossibility, along with necessity and possibility, is often also classified as a modal notion, but it does not call for special discussion here since there is no difficulty in expressing it by the operator $\sim M$ (or alternatively $L\sim$). Propositions which are neither necessary nor impossible are called *contingent*.[2]

A relation between propositions that we may easily express with the tools at our disposal is that of necessary implication. Necessary implication is sometimes called *strict*, in contrast to material, implication, and we shall have more to say about it in Chapter 11. It is important not to confuse $L(p \supset q)$, which means that the whole hypothetical 'if $p$ then $q$' is a necessary truth, or that $q$ follows logically from $p$, with $p \supset Lq$, which means that if $p$ is true then $q$ is a necessary truth. Unhappily, these are often confused in ordinary discourse, sometimes with disastrous results; and neglect of the distinction is made all the easier by the ambiguity of such common idioms as 'If ... then it must be (or is bound to be) the case that —'. To make things worse, the structure of such sentences is more closely analogous to that of $p \supset Lq$, but one suspects that most frequently what the speaker intends to assert (or at least all they are entitled to assert) is something of the form $L(p \supset q)$. Thus someone who says, 'If it rains throughout December it is bound to rain on Christmas Day' probably means to assert that 'it will rain on Christmas Day' follows from 'it will rain throughout December' (which is true, since Christmas Day is in December); but they could be taken to be asserting that if it rains throughout December then it is a necessary truth that it will rain on Christmas Day (which, at least if it does rain throughout December, is false because, come what may about the weather, 'it will rain on Christmas Day' expresses a contingent proposition, not a necessary one).

Perhaps no one, except in their dullest moments, would be taken in by this example. But people have, it appears, confused the necessary truth of 'If a thing is going to happen it is going to happen' with the view that whatever happens happens by logical necessity, or even argued for Fatalism by inferring illicitly from the former to the latter. And in epistemological discussions the fact (if it is a fact) that, of necessity, if someone knows that $p$ then $p$ is true has sometimes been held to show something which does not follow from it at all, viz. that only necessary truths can ever be known. This transition is facilitated if we express the premiss of the argument by the ambiguous but more colloquial 'If you know something, it must be true (can't be false)'. Even a little study of modal logic can protect us from pitfalls in philosophy and elsewhere.

### The language of propositional modal logic

We are now in a position to be able to specify precisely the language we shall use for all the systems of propositional modal logic which we shall describe in later chapters. Its symbols and rules are:

*Primitive symbols*

| | |
|---|---|
| $p, q, r, \ldots$ | [propositional variables] |
| $\sim, L$ | [monadic operators] |
| $\vee$ | [dyadic operator] |
| (, ) | [brackets] |

*Formation rules*

**FR1** A propositional variable is a wff.

**FR2** If $\alpha$ is a wff, so are $\sim\alpha$ and $L\alpha$.





**FR3** If $\alpha$ and $\beta$ are wff, so is $(\alpha \vee \beta)$.

*Definitions*

Def $\wedge$, Def $\supset$, Def $\equiv$ as in PC (p. 6), plus

[Def $M$] $M\alpha =_{\text{df}} \sim L \sim\alpha$

As we did for PC, we adopt the convention that brackets enclosing a complete wff may be omitted.

Clearly every wff of PC is also a wff of modal logic. A few examples of wff of modal logic which are not wff of PC are: $Lp \supset p$; $MLp \supset p$; $L(L(p \vee q) \supset Mq)$; $(Lp \wedge Mq) \supset L(Lp \vee Mq)$; $(MLMp \wedge p) \equiv Lp$.

### Validity in propositional modal logic

Which modal formulae are we to count as valid? It is easy to give a general, intuitive account of validity for modal formulae exactly as we initially did for PC formulae, by saying that a wff is valid iff it 'comes out true' for every uniform replacement of its variables by propositions. In PC, because of the truth-functional nature of all the operators, this initial account led directly to a quite simple formal definition of validity. In modal logic, however, things are not as straightforward; for modal operators are not truth-functional, and it is not at all clear at the outset under what conditions propositions containing them are to count as true or false. The method of defining validity for modal wff which has proved most fruitful and widely applicable is based on the following ideas, which we shall state informally at first but which we shall express more rigorously later on:[3]

(a) Whereas determining the truth-value of a non-modal proposition involves only a consideration of how things actually are, determining the truth-value of a proposition of the form 'Necessarily $p$' or 'Possibly $p$' involves a consideration of how things might have been, of the nature of conceivable states of affairs alternative to the actual one.

(b) For each conceivable state of affairs there is a range of states of affairs which are possible relative to that one. (This reflects the idea we sometimes express by saying that if things were different a new range of possibilities might be opened up, so that things that are not even possible as things stand might be possible then.)

(c) In any given conceivable state of affairs, 'Possibly $p$' counts as true iff $p$ itself would be true in *at least one* state of affairs which is possible relative to that one, and 'Necessarily $p$' counts as true iff $p$ itself would be true in *every* such state of affairs.

With these ideas in mind we shall now describe a more elaborate version of the PC game described on p. 9. We shall call this game the *modal game*. Whereas the PC game involved only one player, in the modal game there can be any number (provided that there is at least one). We are to envisage these players as being seated in some way which determines precisely which players, if any, each player is to be able to see during the course of the game. Screens or some other devices might be used for this purpose; but since in this context being able to see someone means no more than taking note of that person's responses, it will be sufficient to specify, for each player, which players are to be watched and which ignored. There are no restrictions whatsoever on what 'seeing arrangement' among the players may be made: thus we may decide that no one is to be able to see anyone at all, or at the opposite extreme that everyone can see everyone, or we may specify any intermediate arrangement; we may decide that some players shall be able to see themselves while others shall not; if player A can see player B, B may or may not be allowed to see A; and so forth. Finally, before the game begins, each player is provided, as the single player in the PC game was, with a sheet of letters.

We shall call the set of players together with the specification of who is to be able to see whom, a *seating arrangement*, and this together with the players' sheets a *setting* for the modal game, or simply a *setting*.

The game proceeds by calling, to the whole set of players at once, any wff of modal logic we choose, provided that, as in the PC game, its well-formed parts, beginning with the variables, are called first. (We can again assume that the wff are written in primitive notation, with all defined operators eliminated, though we shall, for clarity, state the rule for wff containing $M$ explicitly.)

The instructions for each player are those numbered 1, 2 and 3 in the PC game, together with the following two for calls involving $L$ and $M$:

4. If $L\alpha$ is called (where $\alpha$ is a wff of modal logic), raise your hand if *every player you can see* raised his or her hand when $\alpha$ was called; otherwise keep your hand down.

5. If $M\alpha$ is called, raise your hand if *at least one of the players you can see* raised his or her hand when $\alpha$ was called; otherwise keep your hand down.

As with the PC game, it should be clear that in each setting each wff of modal logic (when appropriately prepared for) will get, from each player, a unique response. In a given setting the call of a formula may of





course lead some players but not others to raise their hands, but if it leads every player without exception to raise his or her hand we shall say that that formula is *successful* in that setting.

How then should we use these games to define validity in propositional modal logic? We have said that our underlying intuitive idea is that a wff should count as valid iff it is true for all values of its variables. In the case of the PC game what this means is that the wff must be successful no matter what sheet is given to the player. Now if we compare the PC game with the modal game, it is not hard to see that the PC game is simply the modal game played in a seating arrangement with just one player and with only PC wff being called. (Strictly speaking there are two possible seating arrangements with one player, according to whether that player can see himself or herself or not; but although these seating arrangements can lead to different results for wff containing $L$ or $M$, they cannot do so for wff of PC.) This suggests that an appropriate generalization of our notion of validity to make it cover modal wff is that of being *valid in a seating arrangement*, in this sense: that a wff $\alpha$ is valid in a given seating arrangement iff in that seating arrangement all the players would raise their hands for $\alpha$, *no matter what sheets were distributed to them* – or, to put this in another way, iff $\alpha$ would be successful in all settings based on that seating arrangement.

If validity is thought of in this way, one consequence is that there will be as many different kinds of validity for modal formulae as there are different seating arrangements, and hence that we can have no unique account of validity in modal logic. At first sight this may seem undesirable; yet on reflection a plurality of criteria of validity is just what our earlier discussion of modal notions would lead us to expect. If 'necessarily' and 'possibly' can be used in a variety of different senses, then it is quite reasonable to suppose that corresponding to each of these senses there will be a different range of acceptable seating arrangements. In fact the possibility of having different kinds of seating arrangements is part of what gives modal logic its richness.

A simple example of a wff which is valid in a certain seating arrangement is $Lp \supset p$. Imagine a seating arrangement in which there are only two players, A and B, and both can see themselves and each other. Take player A. If $p$ is on A's sheet, A will raise his or her hand for $p$, and hence, by the rule for $\supset$, will also raise it for $Lp \supset p$. If $p$ is not on A's sheet, A's hand will not be raised for $p$, and hence, since A can see A, by the rule for $L$ it will not be raised for $Lp$ either. So, by the rule for $\supset$, it must be raised for $Lp \supset p$ in this case also. This means that A must raise his or her hand for $Lp \supset p$, whether $p$ is on A's sheet or not; and B must do likewise, for the same reason.

But although $Lp \supset p$ is valid in this seating arrangement, it is not valid in *every* seating arrangement. For imagine a seating arrangement just like the previous one except that A cannot see himself or herself, and consider a setting in this seating arrangement in which $p$ is on B's list but not on A's. Since B is the only player A can see, A's hand will be raised for $Lp$, but it will not be raised for $p$. So it will not be raised for $Lp \supset p$, and this shows that this wff is not valid in this seating arrangement.

The case of $Lp \supset p$ illustrates some of the richness of modal logic. For it is not difficult to see that this wff is valid not only in the seating arrangement described two paragraphs back, where A and B can see themselves and each other, but also in *any* seating arrangement in which all players can see themselves. And this means that any sense of 'necessary' in which whatever is necessary is true can be reflected by restricting the seating arrangements to those in which all players can at least see themselves.

There are, however, some wff which are valid in *every* seating arrangement. For reasons to be given in the next chapter we shall say that these wff are *K-valid*. It is easy to see that all PC-valid wff are K-valid: for in responding to a PC wff a player in the modal game takes no notice of any other players, and a PC-valid wff is precisely one which any sheet of letters whatsoever would lead a player to raise his or her hand. An example of a specifically modal wff which is K-valid is one which is often called **K**:

**K** $\quad L(p \supset q) \supset (Lp \supset Lq)$

The proof that this wff is K-valid is this: If it were not K-valid then, by the rules for $\supset$, there would have to be a setting in which some player, say A,

(i) raises a hand for $L(p \supset q)$,
(ii) raises a hand for $Lp$,

but

(iii) does not raise a hand for $Lq$.

There cannot, however, be any such setting. For by (iii) there must be a player, say B, whom A can see and whose hand was kept down for $q$. By (ii), since A can see B, B's hand must have been raised for $p$. Hence since B's hand was raised for $p$ but not for $q$, it must have been kept down for $p \supset q$. This, however, conflicts with (i); for since A can





see B, (i) means that B's hand was *raised* for $p \supset q$.

We can think of the modal game in this way: In any setting the players represent conceivable states of affairs or, as they are often called, alternative *possible worlds*, as we spoke of these near the beginning of this section; the players each player is allowed to see represent the states of affairs which are possible relative to the state of affairs which that player represents; and the letters on a player's sheet represent the propositions that are true in that state of affairs. Raising a hand and keeping it down represent respectively truth and falsity in the state of affairs the player represents. Hence what the K-validity of a wff means is that that wff would turn out to be true in every conceivable state of affairs, no matter what propositions we were to replace its variables by, no matter what was true or false in that state of affairs, and no matter what states of affairs were possible relative to that one.

One might at this point raise the question of just what a possible world or conceivable state of affairs really is.[4] This is a matter of some importance and controversy in metaphysics and in the application of modal logic to theories of meaning for natural language. Luckily however, from the point of view of *logic* it makes no difference just what they are, as may be seen from our discussion of the modal game in which the 'worlds' are players. In this book therefore we shall take no position on the ontological status of possible worlds.

## Exercises — 1

1.1 Show that the following wff are valid in every seating arrangement:

(a) $L(p \supset p)$
(b) $(Lp \vee Lq) \supset L(p \vee q)$
(c) $L(p \wedge q) \equiv (Lp \wedge Lq)$
(d) $Mp \supset (Lq \supset Mq)$
(e) $M(p \supset q) \equiv (Lp \supset Mq)$

1.2 Show that in any seating arrangement in which there is a player who cannot see himself or herself $Lp \supset p$ is not valid.

1.3 For each of the following wff devise a seating arrangement in which it is not valid:

(a) $L(p \vee q) \supset (Lp \vee Lq)$
(b) $M(p \supset p)$
(c) $(Lp \supset Lq) \supset L(p \supset q)$
(d) $Lp \supset LLp$

**1.4** (a) Consider a seating arrangement in which every player A can see *at most one* player (who may be A or may be another player). Show that in such a seating arrangement $Mp \supset Lp$ is valid.

(b) Consider a seating arrangement in which a player A can see more than one player. Show that in such a seating arrangement $Mp \supset Lp$ is *not* valid.

**Notes**

[1] Most current logic textbooks give an account of PC in more or less detail. Terminology and notation vary somewhat but this should not confuse the careful reader. Despite its age the fullest introduction to the propositional calculus is still probably found in Church 1956.

[2] The notation $L$ and $M$ for the necessity and possibility operators dates from Feys 1950 (for $L$) and Becker 1930 (for $M$). For a history of notation see appendix 4 of Hughes and Cresswell 1968 (pp. 347–349). The commonly used $\Box$ for $L$ is due to F.B.Fitch and first appears in Barcan 1946. $\Diamond$ for $M$ dates from Lewis and Langford 1932. Other primitives have been studied. Halldén 1949b has a triadic operator in terms of which both the modal operators *and* all the truth-functional operators can be defined. Montgomery and Routley 1966 use contingency $\triangledown$ (or non-contingency, $\vartriangle$) to define the modal operators, though their definitions are only applicable to some systems of modal logic. (See Cresswell 1988.)

[3] The ideas which underlie this account of validity appeared in the late 1950s and early 1960s in the works of Kanger 1957a, Bayart 1958, Kripke 1959 and 1963a, Montague 1960 and Hintikka 1961. Anticipations can be found in Wajsberg 1933, McKinsey 1945, Carnap 1946, Meredith 1956, Thomas 1962 and other works. An algebraic description of this notion of validity is found in Jónsson and Tarski 1951, though the connection with modal logic was not made in that article. Some remarks about the earlier history of modal logic are found in Chapter 11 below.

[4] Some interesting perspectives on this question may be found in the essays in Loux 1979.



# 2

# THE SYSTEMS K, T AND D

## Systems of modal logic

For the rest of Part I we shall be concerned with a number of systems of propositional modal logic. The present chapter will deal with the first three of these. Our way of expounding the systems will be by the axiomatic method. Historically, modal systems were presented in this way before the discovery of an appropriate way to define validity for modal logic, and that is one reason for proceeding as we do. But another, and perhaps more significant, reason is that the axiomatic method allows us to define a class of wff without any reference to their meanings.

An *axiomatic basis* for a logical system consists of (a) a specification of the *language* in which the formulae of the system will be expressed – i.e. a list of *primitive symbols*, together with any definitions that may be thought convenient, together with a set of *formation rules* specifying which strings of symbols are to count as wff; (b) a selected set of wff, known as *axioms*; and (c) a set of *transformation rules*, licensing various operations on the axioms, and also (normally) on wff obtained from the axioms by previous applications of the transformation rules. The wff obtained from the axioms in this way, together with the axioms themselves, are known as the *theorems* of the system. All the systems of propositional modal logic which we shall consider will have the same language, the one specified in the previous chapter on p. 16; so in stating their bases we shall merely list their axioms and transformation rules. An axiomatic basis must be formulated in such a way that we can determine effectively (i) of any arbitrary string of symbols whether or not it is a wff, (ii) of any wff whether or not it is an axiom, and (iii) of any purported application of a transformation rule whether or not it is a genuine application of that rule. We therefore take care that our formulation of formation and transformation rules, and indeed our specification of a system as a whole, can be understood without reference to the interpretation of the symbols; this is often a matter of considerable importance when we come to demonstrate that a system has certain properties. The approach of the last chapter *did*, by contrast, specify a class of formulae: the wff valid in a seating arrangement, in terms of their meaning, for, as we said on p. 20, the players in the games can represent possible worlds, and so the account of validity developed there concerns the relation between symbols and what they stand for. Such an approach is often called a *semantical* approach to logic. An axiomatic approach is then often referred to as a *syntactical* approach.

All this, however, does not mean that in choosing the axioms for a system we ought to keep all thought of interpretation out of our minds. For although we could in theory take any wff whatsoever as axioms, in practice our reason for choosing certain wff as axioms will usually be either that they are valid by some criterion of validity that we have in mind, or at least that they are plausible or interesting in some way which leads us to want to explore their consequences; and these are matters which involve the interpretation we give to our symbols and formulae. Analogously, when we are constructing a system with a certain criterion of validity in mind, we see to it that its transformation rules are such that when they are applied to valid wff the theorems they yield are always valid too. Such transformation rules are said to be *validity-preserving* (with respect to that account of validity).

It is convenient at this point to explain some more of the terminology we shall use in discussing logical systems. When a formula is a theorem of a given system we shall say that it *belongs to*, or is *contained in*, or simply is *in*, that system. If two axiomatic systems, S and S′, have different bases but contain exactly the same theorems, we shall say that S and S′ are *deductively equivalent*, or sometimes simply that they are *equivalent*. If every theorem of S is also a theorem of S′ (whether or not S′ contains other theorems as well) we shall say that S′ *contains* S; thus two systems are deductively equivalent iff each contains the other. If S′ contains all the theorems of S and other theorems as well, we say that it *properly contains* S, or is a *proper extension* of S, and that S′ is the *stronger* and S the *weaker* of the two systems.

## The system K

On p. 20 we introduced the notion of what we called *K-validity*. The first system we shall consider is one which will turn out to have as its theorems precisely those modal formulae which are K-valid. This is usually known nowadays as the system K.[1] Its axioms consist of all valid





wff of PC, i.e. all the wff specified by the following axiom schema,

**PC** If $\alpha$ is a valid wff of PC, then $\alpha$ is an axiom[2]

together with the single distinctively modal wff

**K** $L(p \supset q) \supset (Lp \supset Lq)$

and it has the following three primitive (i.e. initially given) transformation rules:

US (The *Rule of Uniform Substitution*): The result of uniformly replacing any variable or variables $p_1, \dots, p_n$ in a theorem by any wff $\beta_1, \dots, \beta_n$ respectively is itself a theorem.
MP (The *Rule of Modus Ponens*, sometimes also called the *Rule of Detachment*): If $\alpha$ and $\alpha \supset \beta$ are theorems, so is $\beta$.
N (The *Rule of Necessitation*): If $\alpha$ is a theorem, so is $L\alpha$.

Where convenient we shall in future use the following notation:

1. Where $p_1, \dots, p_n$ are some or all of the variables occurring in a wff $\alpha$, and $\beta_1, \dots, \beta_n$ are any wff, we use the expression $\alpha[\beta_1/p_1, \dots, \beta_n/p_n]$ to denote the wff which results from $\alpha$ by replacing $p_1, \dots, p_n$ uniformly by $\beta_1, \dots, \beta_n$ respectively.
2. Where $\alpha$ is a wff and S is an axiomatic system, we write $\vdash_S \alpha$ to mean that that $\alpha$ is a theorem of S. Where no ambiguity is likely to arise we often omit the subscript 'S'.
3. We express the derivability of one wff from one or more other wff by the symbol $\rightarrow$.

Using this notation we could express the transformation rules more succinctly in this way:

US: $\vdash \alpha \rightarrow \ \vdash \alpha[\beta_1/p_1, \dots, \beta_n/p_n]$.
MP: $\vdash \alpha, \alpha \supset \beta \rightarrow \ \vdash \beta$.
N: $\vdash \alpha \rightarrow \ \vdash L\alpha$.

US and MP are not specifically modal rules. US in particular is a rule that it is plausible to require of *any* logical system with a class of symbols to be interpreted as propositional variables, and MP simply reflects the truth-functional meaning of $\supset$. It is easy to see that both these rules are validity-preserving with respect to K-validity, though we shall prove this formally later. N, which *is* a specifically modal rule, also preserves K-validity, for this reason: Suppose $\alpha$ is K-valid – i.e. in every setting every player would raise a hand for $\alpha$; then every player that any player can see would raise a hand for $\alpha$; so by the rule for $L$, every player would raise a hand for $L\alpha$ – i.e. $L\alpha$ is K-valid.

## Proofs of theorems

We have said that the theorems of a system are those wff which can be derived from its axioms by applying its transformation rules. To prove a theorem is therefore to derive it in this way. More precisely, a proof of a theorem $\alpha$ in a system S consists of a finite sequence of wff, each of which is either (i) an axiom of S or (ii) a wff derived from one or more wff occurring earlier in the sequence, by one of the transformation rules or by applying a definition, $\alpha$ itself being the last wff in the sequence. (Note that by this account of what constitutes a proof of a theorem, every wff in a proof is itself a theorem; and also that one reason why we count the axioms themselves as theorems is that any axiom can be thought of as a one-line proof of itself.)

We shall set out proofs in the following way. At the outset we state the theorem to be proved and give it a reference number. Each line of the proof itself contains three items: (a) a wff; (b) a reference number for that wff, written immediately before it; and (c) a justification for writing the wff, written on the left. This justification must consist in showing that the wff satisfies either condition (i) or condition (ii) mentioned above. In case (i) the justification entry consists of the reference number or name of the axiom in question (in the case of an axiom falling under the schema **PC**, if it is listed on p. 13, we cite the name or number assigned to it there; otherwise we simply write '**PC**'). In case (ii) the justification entry refers by number to the earlier wff being used and indicates which transformation rule or definition is being applied. The application of US will be indicated in accordance with the notation explained above, noting within square brackets each variable being replaced and the wff replacing it. The application of MP and N will be indicated by '$\times$ MP' and '$\times$ N' respectively.

We shall first prove two theorems in full detail, and then describe some methods of abbreviating proofs. Theorems will be numbered using the name of the relevant system; thus K1 will be the first theorem we prove in K, and so on.





**K1** $L(p \wedge q) \supset (Lp \wedge Lq)$

PROOF

| | | |
|---|---|---|
| PC1 | (1) | $(p \wedge q) \supset p$ |
| (1) $\times$ N | (2) | $L((p \wedge q) \supset p)$ |
| **K** | (3) | $L(p \supset q) \supset (Lp \supset Lq)$ |
| (3)$[p \wedge q/p, p/q]$ | (4) | $L((p \wedge q) \supset p) \supset (L(p \wedge q) \supset Lp)$ |
| (2), (4) $\times$ MP | (5) | $L(p \wedge q) \supset Lp$ |
| PC2 | (6) | $(p \wedge q) \supset q$ |
| (6) $\times$ N | (7) | $L((p \wedge q) \supset q)$ |
| (3)$[p \wedge q/p]$ | (8) | $L((p \wedge q) \supset q) \supset (L(p \wedge q) \supset Lq)$ |
| (7), (8) $\times$ MP | (9) | $L(p \wedge q) \supset Lq$ |
| PC3 | (10) | $(p \supset q) \supset ((p \supset r) \supset (p \supset (q \wedge r)))$ |
| (10)$[L(p \wedge q)/p, Lp/q, Lq/r]$ | | |
| | (11) | $(L(p \wedge q) \supset Lp) \supset ((L(p \wedge q) \supset Lq) \supset (L(p \wedge q) \supset (Lp \wedge Lq)))$ |
| (5), (11) $\times$ MP | (12) | $(L(p \wedge q) \supset Lq) \supset (L(p \wedge q) \supset (Lp \wedge Lq))$ |
| (9), (12) $\times$ MP | (13) | $(L(p \wedge q) \supset (Lp \wedge Lq))$ **Q.E.D.** |

**K2** $(Lp \wedge Lq) \supset L(p \wedge q)$

PROOF

| | | |
|---|---|---|
| PC4 | (1) | $p \supset (q \supset (p \wedge q))$ |
| (1) $\times$ N | (2) | $L(p \supset (q \supset (p \wedge q)))$ |
| **K** | (3) | $L(p \supset q) \supset (Lp \supset Lq)$ |
| (3)$[q \supset (p \wedge q)/q]$ | | |
| | (4) | $L(p \supset (q \supset (p \wedge q))) \supset Lp \supset L(q \supset (p \wedge q)))$ |
| (2), (4) $\times$ MP | (5) | $Lp \supset L(q \supset (p \wedge q))$ |
| (3)$[q/p, p \wedge q/q]$ | (6) | $L(q \supset (p \wedge q)) \supset (Lq \supset L(p \wedge q))$ |
| PC8 | (7) | $(p \supset q) \supset ((q \supset (r \supset s)) \supset ((p \wedge r) \supset s))$ |
| (7)$[Lp/p, L(q \supset (p \wedge q))/q, Lq/r, L(p \wedge q)/s]$ | | |
| | (8) | $(Lp \supset L(q \supset (p \wedge q))) \supset ((L(q \supset (p \wedge q)) \supset Lq \supset L(p \wedge q))) \supset ((Lp \wedge Lq) \supset L(p \wedge q)))$ |
| (5), (8) $\times$ MP | (9) | $(L(q \supset (p \wedge q)) \supset (Lq \supset L(p \wedge q))) \supset ((Lp \wedge Lq) \supset L(p \wedge q))$ |
| (6), (9) $\times$ MP | (10) | $(Lp \wedge Lq) \supset L(p \wedge q)$ **Q.E.D.** |

The proofs of these theorems satisfy exactly the requirements we listed for a proof in K. Setting out proofs at such length, however, can be not only tedious but sometimes actually a hindrance to understanding the principles which underlie them. We shall therefore introduce a number of conventions which will enable us to state proofs more briefly, while still providing all the information from which a full and rigorously formulated proof could be constructed.

Note first that theorem K2 is the converse of K1. Now we have defined equivalence as mutual implication, so we might expect to be able to use K1 and K2 to obtain $L(p \wedge q) \equiv (Lp \wedge Lq)$ as a new theorem. And in fact PC5 will enable us to do this; for if we substitute $L(p \wedge q)$ for $p$ and $(Lp \wedge Lq)$ for $q$ in PC5, and then apply MP twice, using K1 the first time and K2 the second time, the result will be precisely $L(p \wedge q) \equiv (Lp \wedge Lq)$. How shall we set all this out as a proof? If we are to adhere strictly to our criteria for a proof, we cannot use K1 (or K2) until we have written it down, and we are not allowed to write it down until we have derived it from axioms and earlier wff in the sequence which forms the proof; but this means that our proof of our new theorem will have to incorporate complete proofs of K1 and K2 before we begin to use these theorems in combination with PC5. Setting out the proof like this, however, involves a quite wasteful repetition of work that we have already done in proving K1 and K2 themselves. We shall therefore adopt the convention that *after* we have proved any theorem, we may write that theorem as a line in any subsequent proof, simply citing its reference number as its justification. The proof of our new theorem will then look like this:

**K3** $L(p \wedge q) \equiv (Lp \wedge Lq)$

PROOF

| | | |
|---|---|---|
| K1 | (1) | $L(p \wedge q) \supset (Lp \wedge Lq)$ |
| K2 | (2) | $(Lp \wedge Lq) \supset L(p \wedge q)$ |
| PC5 | (3) | $(p \supset q) \supset ((q \supset p) \supset (p \equiv q))$ |
| $(3)[L(p \wedge q)/p, Lp \wedge Lq/q]$ | (4) | $(L(p \wedge q) \supset (Lp \wedge Lq)) \supset (((Lp \wedge Lq) \supset L(p \wedge q)) \supset (L(p \wedge q) \equiv (Lp \wedge Lq)))$ |
| (1), (4) × MP | (5) | $((Lp \wedge Lq) \supset L(p \wedge q)) \supset (L(p \wedge q) \equiv (Lp \wedge Lq))$ |
| (2), (5) × MP | (6) | $L(p \wedge q) \equiv (Lp \wedge Lq)$ **Q.E.D.** |

K3 may be called the *Law of L-distribution.*





Consider next how we used PC5 in the above proof. What we did was to make substitutions in it which produced, at line (4), an implicative wff whose antecedent was an already proved wff (K1) and whose consequent had as *its* antecedent another already proved wff (K2). We then used MP twice to obtain the consequent of its consequent as a theorem. Now it should be clear that we can use PC5 in this way not only in the case of K1 and K2, but whenever we have already proved both a wff of the form $\alpha \supset \beta$ and its converse $\beta \supset \alpha$; i.e. by substituting $\alpha$ for $p$ and $\beta$ for $q$ in PC5 and applying MP twice, we can obtain $\alpha \equiv \beta$. We thus have a rule which could be expressed in this way:

$$\vdash \alpha \supset \beta,\ \vdash \beta \supset \alpha \rightarrow\ \vdash \alpha \equiv \beta$$

This rule is not part of the axiomatic basis of K. Nevertheless it is what we call a *derived* rule of K, in the sense that we may always use it as a transformation rule in a proof, since anything we can prove by using it we could also prove, though at greater length, from the axiomatic basis alone. To establish that a rule is a derived rule of a system we simply show how we could always do without it. In the present case we can do this as follows:

| | | |
|---|---|---|
| Given: | (1) | $\alpha \supset \beta$ |
| Given: | (2) | $\beta \supset \alpha$ |
| PC5 | (3) | $(p \supset q) \supset ((q \supset p) \supset (p \equiv q))$ |
| $(3)[\alpha/p, \beta/q]$ | (4) | $(\alpha \supset \beta) \supset ((\beta \supset \alpha) \supset (\alpha \equiv \beta))$ |
| (1), (4) × MP | (5) | $(\beta \supset \alpha) \supset (\alpha \equiv \beta)$ |
| (2), (5) × MP | (6) | $\alpha \equiv \beta$ **Q.E.D.** |

Since all we have used in establishing this rule (apart from US and MP) is PC5, we shall signal its use in justification entries simply by writing '× PC5'.

The procedure we have described for the use of PC5 will in fact enable us to derive a rule of K from any valid PC wff whose main operator is $\supset$. For if $\alpha$ is a valid PC wff, it is an axiom of K, and hence, by US, all its substitution-instances are theorems of K. So if we can make substitutions for the variables in $\alpha$ which will turn it into a wff whose antecedent is a wff we have already proved, we can use MP to detach its consequent and count that as a theorem too. (This is why MP is sometimes called *Detachment.*) In cases such as PC5 itself where the PC axiom has the overall form A $\supset$ (B $\supset$ C), if we can make substitutions which will turn both A and B into already proved wff, we can then use MP twice to obtain the result of these substitutions in C. A specially useful PC axiom of this kind is PC6, to which we gave the name *Syll* on p. 13. This gives us the rule

$$\vdash \alpha \supset \beta,\ \vdash \beta \supset \gamma \rightarrow\ \vdash \alpha \supset \gamma$$

which says that when we have proved two implicative wff in which the consequent of one is the antecedent of the other, we can count as a theorem the implicative wff whose antecedent is the antecedent of the former and whose consequent is the consequent of the latter. We shall indicate the application of this rule by '× Syll', and give analogous indications, by name or number, of other rules similarly derived from PC axioms. In cases where the PC wff is not one we have listed we shall write simply × PC.

Another way of shortening the statement of proofs is this. Line (3) in the proof of K1 is simply the axiom **K** itself, and line (4) is derived from this by US. The presence of **K** (without substitutions) is required by our definition of what counts as a proof; but it would be more economical, and still give all the information from which a detailed proof could be constructed, to omit line (3) altogether and give **K** with the appropriate substitutions as the justification for line (4). Similarly, we could omit line (10) and give PC3 with the appropriate substitutions as the justification for line (11). Somewhat analogously, we could omit line (1) and give 'PC1 × N' as the justification for immediately writing the present line (2). So we shall adopt the convention that citing any axioms or previously proved theorems by name or number and indicating the application of a transformation rule to them will be a sufficient justification entry for the wff obtained thereby.

Finally, by using **K** together with N, US and MP, we can obtain a very useful derived rule. This is a specifically modal rule and we shall give it a special name as the first such rule we shall prove:

**DR1** $\vdash \alpha \supset \beta \rightarrow\ \vdash L\alpha \supset L\beta$

PROOF

| | | | |
|---|---|---|---|
| Given: | (1) | $\alpha \supset \beta$ | |
| (1) × N | (2) | $L(\alpha \supset \beta)$ | |
| K[$\alpha/p$, $\beta/q$] | (3) | $L(\alpha \supset \beta) \supset (L\alpha \supset L\beta)$ | |
| (2), (3) × MP | (4) | $L\alpha \supset L\beta$ | Q.E.D. |





In the light of all this let us see how we can set out the proofs of K1–K3 in the abbreviated style which we shall use from now on:

**K1** $L(p \wedge q) \supset (Lp \wedge Lq)$

PROOF

| | | | |
|---|---|---|---|
| PC1 × DR1 | (1) | $L(p \wedge q) \supset Lp$ | |
| PC2 × DR1 | (2) | $L(p \wedge q) \supset Lq$ | |
| (1), (2) × PC3 | (3) | $L(p \wedge q) \supset (Lp \wedge Lq)$ | Q.E.D. |

**K2** $(Lp \wedge Lq) \supset L(p \wedge q)$

PROOF

| | | | |
|---|---|---|---|
| PC4 × DR1 | (1) | $Lp \supset L(q \supset (p \wedge q))$ | |
| K[$q/p$, $p \wedge q/q$] | (2) | $L(q \supset (p \wedge q)) \supset (Lq \supset L(p \wedge q))$ | |
| (1), (2) × PC8 | (3) | $(Lp \wedge Lq) \supset L(p \wedge q)$ | Q.E.D. |

**K3** $L(p \wedge q) \equiv (Lp \wedge Lq)$

PROOF

| | | | |
|---|---|---|---|
| K1, K2 × PC5 | (1) | $L(p \wedge q) \equiv (Lp \wedge Lq)$ | Q.E.D. |

We shall now prove some more theorems and derived rules of K.

**K4** $(Lp \vee Lq) \supset L(p \vee q)$

PROOF

| | | | |
|---|---|---|---|
| PC9 × DR1 | (1) | $Lp \supset L(p \vee q)$ | |
| PC10 × DR1 | (2) | $Lq \supset L(p \vee q)$ | |
| (1), (2) × PC11 | (3) | $(Lp \vee Lq) \supset L(p \vee q)$ | Q.E.D. |

Note that K4, unlike K3, is only an implication, not an equivalence. The converse of K4 is not a theorem of **K**, and in fact at the intuitive level is not a valid formula: it may be necessary that you are awake or asleep without its being necessary that you are awake or its being necessary that you are asleep.

We next prove two further derived rules. The first of these is:

**DR2** $\vdash \alpha \equiv \beta \rightarrow\ \vdash L\alpha \equiv L\beta$

PROOF

| | | |
|---|---|---|
| Given: | (1) | $\alpha \equiv \beta$ |
| (1) × PC | (2) | $\alpha \supset \beta$ |
| (2) × DR1 | (3) | $L\alpha \supset L\beta$ |
| (1) × PC | (4) | $\beta \supset \alpha$ |
| (4) × DR1 | (5) | $L\beta \supset L\alpha$ |
| (3), (5) × PC5 | (6) | $L\alpha \equiv L\beta$ |

Q.E.D.

Note that in this proof we used purely PC principles to get from $\alpha \equiv \beta$ at line (1) to both $\alpha \supset \beta$ and $\beta \supset \alpha$ at lines (3) and (5). Clearly we could do this with any theorem which has the form of an equivalence, and for this reason whenever we have proved a wff of the form $\alpha \equiv \beta$ we shall assume that we have proved both $\alpha \supset \beta$ and $\beta \supset \alpha$; for example, if we have proved $\alpha \equiv \beta$ and $\alpha$, we shall assume that $\beta$ follows, and if we have proved $\alpha \equiv \beta$ and $\beta$, we shall assume that $\alpha$ follows, by MP in each case.

Our next derived rule is that of *Substitution of Equivalents*, which we shall usually call *Eq*. What this states is that if $\alpha$ is a theorem and $\beta$ differs from $\alpha$ only in having some wff, $\delta$, at one or more places where $\alpha$ has a wff, $\gamma$, then if $\gamma \equiv \delta$ is a theorem, $\beta$ is a theorem. In other words, if we have proved $\gamma \equiv \delta$, we can replace $\gamma$ by $\delta$ in any theorem (not necessarily uniformly), and the result will also be a theorem. We now want to show that this rule holds in K. To do so we first note that the following are valid wff of PC, and therefore axioms of K:

$$(p \equiv q) \supset (\sim p \equiv \sim q)$$
$$(p \equiv q) \supset ((p \vee r) \equiv (q \vee r))$$
$$(p \equiv q) \supset ((r \vee p) \equiv (r \vee q))$$

Suppose now that $\gamma \equiv \delta$ is a theorem of K. Then by substitution in these three axioms, and MP, it follows that the following are also theorems of K,

$$\sim\gamma \equiv \sim\delta$$
$$(\gamma \vee \zeta) \equiv (\delta \vee \zeta)$$
$$(\zeta \vee \gamma) \equiv (\zeta \vee \delta)$$

for any wff $\zeta$. DR2, which we proved above, enables us to add to this list of consequences of $\gamma \equiv \delta$, $L\gamma \equiv L\delta$.

From this it follows that if $\alpha$ is any wff which is built up from $\gamma$ using $\sim$ and $L$ as the only monadic operators and $\vee$ as the only dyadic one, and $\beta$ is built up from $\delta$ in exactly the same way as $\alpha$ is from $\gamma$, then if $\gamma \equiv \delta$ is a theorem, so is $\alpha \equiv \beta$; and therefore, if $\alpha$ is a theorem, then by MP so is $\beta$. Since every modal wff can be written with $\sim$, $L$ and $\vee$ as its only operators, what we have just shown is that we can apply Eq unrestrictedly in K; i.e. whenever we have a theorem of K of the form $\gamma \equiv \delta$, we can replace $\gamma$ by $\delta$ in any theorem $\alpha$, no matter where $\gamma$ occurs in $\alpha$, and the result will also be a theorem of K.

Where an equivalential wff has a name, e.g. K3, and we are using Eq to replace an instance of one side of the equivalence by an instance of the other side in some wff, we shall indicate the application of Eq by (in this example) '× K3 × Eq', and analogously in other cases. A rich source of equivalential wff is of course provided by valid PC equivalences.

## *L* and *M*

Our next theorem, which will help us to establish another extremely useful derived rule, is:

**K5** $Lp \equiv \sim M\sim p$

PROOF

| | | |
|---|---|---|
| PC12 (DN) | (1) | $p \equiv \sim\sim p$ |
| (1)[$Lp/p$] | (2) | $Lp \equiv \sim\sim Lp$ |
| (2) × (1) × Eq: | (3) | $Lp \equiv \sim\sim L\sim\sim p$ |
| (3) Def $M$ | (4) | $Lp \equiv \sim M\sim p$ |

Q.E.D.

Clearly K5, by Eq, will entitle us to replace $L$ by $\sim M\sim$ anywhere in a theorem; and by Def $M$ we may replace $M$ anywhere in a theorem by $\sim L\sim$. (By saying that we are 'entitled' to do these things, or 'may' do them we simply mean that the result of doing them is itself a theorem.) The rule we are about to state is a kind of generalization of these procedures. We shall call it the *Rule of L–M Interchange* ('LMI' for short), and what it states is that in any sequence of adjacent monadic modal operators ($L$s and $M$s) in a theorem, $L$ may be replaced by $M$ and $M$ by $L$ throughout, provided that a $\sim$ is either inserted or deleted both immediately before and immediately after the sequence. (Thus $LM$ may be replaced by $\sim ML\sim$, $\sim LLL$ by $MMM\sim$, $MLLM\sim$ by $\sim LMML$, and so forth.)

We shall now establish that this rule holds in K. Let $A_1 \ldots A_n$ be a sequence of monadic modal operators (i.e. each $A_i$ is either $L$ or $M$). For each $A_i$, let $A_i'$ be $M$ if $A_i$ is $L$, and $L$ if $A_i$ is $M$. We first show that

$$(*) \quad A_1 \ldots A_n p \equiv \sim A_1' \ldots A_n' \sim p$$

is a theorem of K. To do so we begin with the following substitution-instance of the PC valid wff $p \equiv p$:

$$(1) \quad A_1 \ldots A_n p \equiv A_1 \ldots A_n p$$

Next, in the right-hand side of (1) we replace each $M$ by $\sim L \sim$ (by Def $M$) and each $L$ by $\sim M \sim$ (by K5 and Eq). The result will be:

$$(2) \quad A_1 \ldots A_n p \equiv \sim A_1' \sim \sim A_2' \sim \ldots \sim A_{n-1}' \sim \sim A_n' \sim p$$

We now use DN ($p \equiv \sim\sim p$) and Eq to delete all occurrences of $\sim\sim$ in (2), and the result is (*) as required. Appropriate substitutions for $p$ in (*), and Eq, will then entitle us to replace any sequence $A_1 \ldots A_n$ by $\sim A_1' \ldots A_n' \sim$ in any theorem. Finally, if the sequence before replacement was immediately preceded or followed by $\sim$, the result of the replacement will give us $\sim\sim$ at the beginning or the end of the new sequence, and this may be deleted by DN and Eq. We have thus shown that every application of LMI to a theorem of K results in a theorem of K – i.e. we have established LMI as a derived rule of K.

Note that the sequence to which we apply LMI may have only a single member. Applications of K5 and Def $M$ are thus themselves applications of LMI, and when convenient we shall indicate them too by '× LMI'. Note too that there is nothing to prevent us applying LMI only to part of a sequence; e.g. we may apply LMI to the first three operators in $LMMLM$, leaving the last two unaltered, and thus obtain $\sim MLL \sim LM$.

**K6** $M(p \vee q) \equiv (Mp \vee Mq)$

PROOF

| | | |
|---|---|---|
| K3 [$\sim p/p, \sim q/q$] | (1) | $L(\sim p \wedge \sim q) \equiv (L\sim p \wedge L\sim q)$ |
| (1) × LMI | (2) | $\sim M \sim(\sim p \wedge \sim q) \equiv (\sim Mp \wedge \sim Mq)$ |
| (2) × PC13 × Eq | (3) | $\sim M(p \vee q) \equiv (\sim Mp \wedge \sim Mq)$ |
| (3) × PC | (4) | $M(p \vee q) \equiv (Mp \vee Mq)$ **Q.E.D.** |

K6 expresses the same kind of principle for possibility and disjunction as K3 does for necessity and conjunction; it may be called the *Law of M-*





*distribution.*

**K7** $M(p \supset q) \equiv (Lp \supset Mq)$

PROOF

| | | |
|---|---|---|
| K6[$\sim p/p$] | (1) | $M(\sim p \vee q) \equiv (M\sim p \vee Mq)$ |
| (1) × LMI | (2) | $M(\sim p \vee q) \equiv (\sim Lp \vee Mq)$ |
| (2) Def $\supset$ | (3) | $M(p \supset q) \equiv (Lp \supset Mq)$ **Q.E.D.** |

We now derive a rule which is like DR1 except that $M$ takes the place of $L$.

**DR3** $\vdash \alpha \supset \beta \rightarrow \vdash M\alpha \supset M\beta$

PROOF

| | | |
|---|---|---|
| Given: | (1) | $\alpha \supset \beta$ |
| (1) × PC15(Transp) | (2) | $\sim\beta \supset \sim\alpha$ |
| (2) × DR1 | (3) | $L\sim\beta \supset L\sim\alpha$ |
| (3) × PC | (4) | $\sim L\sim\alpha \supset \sim L\sim\beta$ |
| (4) Def $M$ | (5) | $M\alpha \supset M\beta$ **Q.E.D.** |

Note that by repeated applications of DR1 and/or DR3 we can prefix any sequence of modal operators to both sides of an implicative theorem.

**K8** $M(p \wedge q) \supset (Mp \wedge Mq)$

We shall give two ways of proving K8. The first uses DR3 in the way that the proof of K4 used DR1, and the second obtains K8 from K4 in the same manner as K6 was obtained from K3. Here is the first:

PROOF

| | | |
|---|---|---|
| PC1 × DR3 | (1) | $M(p \wedge q) \supset Mp$ |
| PC2 × DR3 | (2) | $M(p \wedge q) \supset Mq$ |
| (1), (2) × PC3 | (3) | $M(p \wedge q) \supset (Mp \wedge Mq)$ **Q.E.D.** |

Here is the second proof:

PROOF

| | | |
|---|---|---|
| K4[$\sim p/p, \sim q/q$] | (1) | $(L\sim p \vee L\sim q) \supset L(p \vee q)$ |
| (1) × PC15 × Eq | (2) | $\sim L(\sim p \vee \sim q) \supset \sim(L\sim p \vee L\sim q)$ |

| | | |
|---|---|---|
| (2) × LMI | (3) $M{\sim}({\sim}p \vee {\sim}q) \supset {\sim}({\sim}Mp \vee {\sim}Mq)$ | |
| (3) × PC14 × Eq | (4) $M(p \wedge q) \supset (Mp \wedge Mq)$ | **Q.E.D.** |

As was the case with K4, but in contrast with K6, the converse of K8 is not a theorem of K. We do however have the following partial converse to K4:

**K9** $L(p \vee q) \supset (Lp \vee Mq)$

PROOF

| | | |
|---|---|---|
| K[$\sim q/p, p/q$] | (1) $L({\sim}q \supset p) \supset (L{\sim}q \supset Lp)$ | |
| (1) Def $\supset$, × DN(PC12) | (2) $L(q \vee p) \supset ({\sim}L{\sim}q \vee Lp)$ | |
| (2) Def $M$, × Comm(PC16) | (3) $L(p \vee q) \supset (Lp \vee Mq)$ | **Q.E.D.** |

**Validity and soundness**

As we remarked earlier in this chapter, the theorems of the system K will turn out to be precisely those wff which are K-valid in the sense explained on p. 20. It is important to be quite clear that this is a substantive fact, and not something which is true by definition, as our use of the label 'K-valid' might at first suggest. To be a theorem of K is to be derivable from the axioms of K by the transformation rules of K; to be K-valid is to be successful in every setting of the modal game. We have here two distinct concepts, and the fact that a wff is a theorem of K iff it is K-valid is something we have to *prove*, not something we can assume. We shall in fact come across many cases in which we have an axiomatic modal system defined without any reference to an account of validity, and a definition of validity formulated without any reference to theoremhood in a system, and yet the theorems of that system are precisely the wff which are valid by that definition; but this is something which has to be proved in every case, and it should be obvious that giving the system and the validity-definition the same name (as we shall often do) does nothing to prove it but serves to remind us of the connection once it has been proved. To show that there is a match of this kind between a system and a validity definition we have to prove two things: (A) that every theorem of the system is valid by that definition, and (B) that every wff valid by that definition is a theorem of the system. If (A) holds, we say that the system is *sound*, and if (B) holds we say that it is *complete*, in each case with respect to the validity-definition in question. The completeness of a system is usually more difficult to establish than its soundness, and we shall defer the task of proving the completeness of K till Chapter 6. Here, however, we shall give a proof of its soundness with respect to K-validity.





In a sense we have done this already; for on p. 20 we gave a proof that all valid wff of PC and the wff **K** (i.e. all the axioms of K) are K-valid, and earlier in the present chapter we at least sketched an argument to show that the transformation rules of K preserve K-validity. We shall now, however, give a more rigorous definition of validity for modal formulae and in terms of it a more formally exact proof of the soundness of K.

Our account of the modal game on p. 18, though it was intended to make the idea of validity more immediately comprehensible, had both certain inessential features and also certain limitations, which we now want to remove. It ought not to be difficult to see that speaking of players at all, of some players being able or unable to see other players, and of the raising or non-raising of hands, is quite inessential to the logical structure of the test that is being applied to formulae. Instead of a set of human players we could have a collection of objects of any kind at all; but to reflect the idea, mentioned on p. 21, that the players represent alternative ways the world might be, these objects are sometimes called 'possible worlds', or simply 'worlds', and this is the terminology that we shall usually employ in this book. Similarly, it does not matter what takes the place of the seeing-relation among the players, so long as it is some kind of dyadic relation, R, defined over the objects in question, in the sense that it is specified for every pair of these objects, $w$ and $w'$, whether or not $w\mathrm{R}w'$. Sometimes R is called the *accessibility-relation*, and when $w\mathrm{R}w'$, $w'$ is said to be *accessible from w*, or to be *possible relative to w*. (In this book we shall sometimes use this terminology, but we shall also, when convenient, carry over a metaphor derived from the modal game and speak of one world being able to *see* another. The point to be clear about is that, whatever terminology we use, from a formal point of view R is no more than a relation which may or may not hold between any pair of worlds.)

In describing the modal game we called a set of players and a specification of which players could see which a *seating arrangement*. In our present more abstract account we call the pair $\langle W,R \rangle$, where W is a set of worlds and R is a specification of which of these is related to which, a *frame*.[3] We note here one limitation involved in our description of the modal game, which we can now remove. In any 'real life' attempt to play the modal game, the number of players involved would have to be finite, and in fact in practice fairly small; but we need place no limits to the number of worlds in a frame – there may be only one, there may be 17, there may be infinitely many.

Within each seating arrangement in the modal game we could have any number of *settings* by giving each player a list of variables. As we also remarked on p. 21, this corresponds to the idea that those variables are true, or are assigned the value 1, in the state of affairs represented by the player in question, with the other variable being false, or assigned the value 0. Again, there is a limitation here if we take the game literally, since in practice any list of variables would have to be finite; but we do not wish to have any such restriction in the formal definition of validity which we are now constructing. We shall refer to an assignment of values within a frame as V, and where $p$ is any propositional variable and $w$ is any world in the frame (i.e. $w \in$ W),[4] we shall write $\mathrm{V}(p,w) = 1$ if V assigns the value 1 to $p$ in $w$, and $\mathrm{V}(p,w) = 0$ if it assigns the value 0 to it. Where $\langle \mathrm{W,R} \rangle$ is a frame and V is a value-assignment within that frame, we call $\langle \mathrm{W,R,V} \rangle$ a *model*, and more specifically a model *based on* the frame $\langle \mathrm{W,R} \rangle$. Thus a model corresponds to a setting in the modal game.

We can set out all this as follows:

A *frame* is an ordered pair $\langle \mathrm{W,R} \rangle$, where W is a non-empty set of objects (worlds), and R is a dyadic relation defined over the members of W, i.e. it is determinate for any (not necessarily distinct) $w$ and $w'$ in W whether or not $w\mathrm{R}w'$.

A *model* is an ordered triple $\langle \mathrm{W,R,V} \rangle$ where $\langle \mathrm{W,R} \rangle$ is a frame and V is a value-assignment satisfying the following conditions:

1. For any propositional variable, $p$, and any $w \in$ W, either $\mathrm{V}(p,w) = 1$ or $\mathrm{V}(p,w) = 0$.
2. [V$\sim$] For any wff, $\alpha$, and any $w \in$ W, $\mathrm{V}(\sim\alpha,w) = 1$ if $\mathrm{V}(\alpha,w) = 0$; otherwise $\mathrm{V}(\sim\alpha,w) = 0$.
3. [V$\vee$] For any wff $\alpha$ and $\beta$, and for any $w \in$ W, $\mathrm{V}((\alpha \vee \beta),w) = 1$ if either $\mathrm{V}(\alpha,w) = 1$ or $\mathrm{V}(\beta,w) = 1$; otherwise $\mathrm{V}((\alpha \vee \beta),w) = 0$.
4. [V*L*] For any wff $\alpha$ and for any $w \in$ W, $\mathrm{V}(L\alpha,w) = 1$ if for every $w' \in$ W such that $w\mathrm{R}w'$, $\mathrm{V}(\alpha,w') = 1$; otherwise $\mathrm{V}(L\alpha,w) = 0$.

Although the conditions for the other operators we have introduced are strictly unnecessary, since all wff can be written in primitive notation, we give them here for ease of reference:

[V$\wedge$] For any wff $\alpha$ and $\beta$, and for any $w \in$ W, $\mathrm{V}((\alpha \wedge \beta),w) = 1$ if both $\mathrm{V}(\alpha,w) = 1$ and $\mathrm{V}(\beta,w) = 1$; otherwise $\mathrm{V}((\alpha \wedge \beta),w) = 0$.

[V$\supset$] For any wff $\alpha$ and $\beta$, and for any $w \in$ W, $\mathrm{V}((\alpha \supset \beta),w) = 1$ if either $\mathrm{V}(\alpha,w) = 0$ or $\mathrm{V}(\beta,w) = 1$; otherwise $\mathrm{V}((\alpha \supset \beta),w) = 0$.





[V$\equiv$] For any wff $\alpha$ and $\beta$, and for any $w \in$ W, $\mathrm{V}((\alpha \equiv \beta),w) = 1$ if $\mathrm{V}(\alpha,w) = \mathrm{V}(\beta,w)$; otherwise $\mathrm{V}((\alpha \equiv \beta),w) = 0$.

[V*M*] For any wff $\alpha$ and for any $w \in$ W, $\mathrm{V}(M\alpha,w) = 1$ if for some $w' \in$ W such that $w\mathrm{R}w'$, $\mathrm{V}(\alpha,w') = 1$; otherwise $\mathrm{V}(M\alpha,w) = 0$.

A model $\langle \mathrm{W,R,V} \rangle$ is said to be *based on* the frame $\langle \mathrm{W,R} \rangle$.

We now define validity on a frame by saying that a wff $\alpha$ is *valid on* a frame $\langle \mathrm{W,R} \rangle$ iff, for every model $\langle \mathrm{W,R,V} \rangle$ based on $\langle \mathrm{W,R} \rangle$, and for every $w \in$ W, $\mathrm{V}(\alpha,w) = 1$. Finally we define K-validity by saying that a wff is *K-valid* iff it is valid on every frame.

We are now in a position to prove the soundness of K with respect to K-validity as we have just defined this. Our method of doing so will in fact yield a more general result which we shall be able to use to prove the soundness of many other systems.

THEOREM 2.1 Every theorem of K is K-valid.[5]

What we have to prove is that every wff derivable from the axioms of K by the transformation rules of K is valid on every frame. For this it is clearly sufficient to prove (1) that every axiom of K is valid on every frame, and (2) that the rules US, MP and N preserve validity on a frame – i.e. that if they are applied to wff which are valid on any given frame, the resulting wff are also valid on that frame. In stating the more general consequence which we mentioned above we shall use the following terminology: where $\Lambda$ is any set of modal wff (which may have only one member or more than one – even infinitely many members), we let 'K + $\Lambda$' denote the axiomatic system obtained by adding to K, as extra axioms, all the wff in $\Lambda$ (and retaining the transformation rules US, MP and N).[6] Our more general result is this:

THEOREM 2.2 If $\Lambda$ is any set of modal wff and $\langle \mathrm{W,R} \rangle$ is a frame on which each wff in $\Lambda$ is valid, then every theorem of K + $\Lambda$ is valid on $\langle \mathrm{W,R} \rangle$.

As we have noted, the soundness of K with respect to K-validity (theorem 2.1) follows immediately from theorem 2.2. Theorem 2.2 follows from the following two lemmas:

LEMMA 2.3 If $\langle \mathrm{W,R} \rangle$ is any frame, every valid PC wff is valid on $\langle \mathrm{W,R} \rangle$, and so is the wff K.

LEMMA 2.4 Where $\langle W,R\rangle$ is any frame,

(i) if $\alpha$ is valid on $\langle W,R\rangle$, so is $\alpha[\beta_1/p_1, \dots ,\beta_n/p_n]$ (i.e. $\alpha$ with $\beta_1, \dots , \beta_n$ uniformly replacing $p_1, \dots ,p_n$ respectively);
(ii) if $\alpha$ and $\alpha \supset \beta$ are both valid on $\langle W,R\rangle$, so is $\beta$;
(iii) if $\alpha$ is valid on $\langle W,R\rangle$, so is $L\alpha$.

We shall prove the lemmas in a moment, but before doing so we shall note that theorem 2.2 is an immediate consequence of lemmas 2.3 and 2.4, since by lemma 2.3 every axiom of K is valid on *every* frame, and by lemma 2.4 the transformation rules preserve validity on any frame whatsoever. The importance of theorem 2.2 can be indicated in this way. Apart from a few systems which we shall mention in Chapters 11 and 12, and which stand a little outside mainstream modal logic, K is the weakest of the modal systems we shall be discussing. Each of the other systems will be a proper extension of K (i.e. it will contain not only all the theorems of K but other theorems as well). Modal systems which contain K (including K itself) together with US, MP and N are commonly known as *normal* modal systems, and we shall usually present these other systems by adding one or more extra axioms to the basis of K. For each such system we shall also have (or at least we shall try to find) a definition of validity which matches it in the way that K-validity matches the system K; i.e., which is such that the theorems of the system are precisely the wff which are valid by that definition. Typically we shall produce such a definition by specifying a certain class $\mathscr{C}$ of frames, and saying that a wff is valid with respect to $\mathscr{C}$ ($\mathscr{C}$-valid) iff it is valid on every frame in $\mathscr{C}$. And when a system S is both sound and complete with respect to a class $\mathscr{C}$ of frames, so that the theorems of S consist of all and only those wff that are valid on every frame in $\mathscr{C}$, we say that S is *characterized by* $\mathscr{C}$. To come at last to the importance of theorem 2.2: what it tells us is that if we have a system K + $\Lambda$ and a class of frames $\mathscr{C}$, then in order to prove that K + $\Lambda$ is sound with respect to $\mathscr{C}$, all we have to do is to show that every wff in $\Lambda$ is valid on every frame in $\mathscr{C}$.

We note here some of the terminology we shall use in discussing frames and models. If every theorem of a system S is valid on a frame $\langle W,R\rangle$, we say that $\langle W,R\rangle$ is a *frame for* S. If a wff $\alpha$ is not valid on a given frame we sometimes say that it *fails* on that frame, or that it can be *falsified* on that frame. A model in which $\alpha$ is false in at least one world is called a *falsifying model* for $\alpha$.

So now what remains is to prove lemmas 2.3 and 2.4.





*Proof of lemma 2.3*: (A) In any model, a PC wff is evaluated in any world without reference to any other world. Therefore, since a valid PC wff has the value 1 for every value-assignment to the variables, it has the value 1 in every world in every model, i.e. it is valid on every frame. (B) If K were not valid on every frame, there would have to be a model $\langle W,R,V\rangle$ in which for some $w \in W$, (i) $V(L(p \supset q),w) = 1$, (ii) $V(Lp,w) = 1$, and (iii) $V(Lq,w) = 0$. There cannot, however, be any such model. For by (iii), there must be some $w' \in W$ such that $wRw'$ and $V(q,w') = 0$; by (ii), since $wRw'$, $V(p,w') = 1$; hence, by [V$\supset$], $V((p \supset q),w') = 0$; but then by [V*L*], since $wRw'$, we have $V(L(p \supset q),w) = 0$, which contradicts (i).

*Proof of lemma 2.4*:
(i) Suppose that $\langle W,R\rangle$ is a frame and $\alpha[\beta_1/p_1, \dots ,\beta_n/p_n]$ is not valid on $\langle W,R\rangle$. Then there is a model $\langle W,R,V\rangle$ based on $\langle W,R\rangle$ such that for some $w^* \in W$, $V(\alpha[\beta_1/p_1, \dots ,\beta_n/p_n],w^*) = 0$. Let $\langle W,R,V^*\rangle$ be a model based on the same frame $\langle W,R\rangle$, in which $V^*$ is just like V except that for any $w \in W$, and any $1 \leq i \leq n$, $V^*(p_i,w) = V(\beta_i,w)$. Then $V^*(\alpha,w^*) = 0$, and so $\alpha$ is not valid on $\langle W,R\rangle$. (What this amounts to is simply that whatever model falsifies $\alpha[\beta_1/p_1, \dots ,\beta_n/p_n]$, if we had given the variables that have been replaced the same values as the wff that have replaced them, then we could have falsified the original $\alpha$, showing that it wasn't valid in the first place.)
(ii) If both $\alpha$ and $\alpha \supset \beta$ are valid on $\langle W,R\rangle$, then in every world in every model based on $\langle W,R\rangle$, both $\alpha$ and $\alpha \supset \beta$ are true; hence by [V$\supset$] so is $\beta$; i.e., $\beta$ is valid on $\langle W,R\rangle$.
(iii) If $\alpha$ is valid on $\langle W,R\rangle$, then in every world in every model based on $\langle W,R\rangle$, $\alpha$ is true; hence for every such world, $\alpha$ is true in every world which it can see; so $L\alpha$ is true in every such world – i.e., $L\alpha$ is valid on $\langle W,R\rangle$.

## The system T

On p. 20 we showed that the wff $Lp \supset p$ is not K-valid. In the light of theorem 2.1, this means that it is not a theorem of K. We could, however, add it as an extra axiom to obtain a system stronger than K itself. Now what the formula means is that whatever is necessarily so is so, and we remarked on p. 14 that although there are some senses of 'necessarily' for which this does not hold, there are others for which it does; we therefore have a motive for constructing a system or systems which will reflect these latter senses. The system obtained by adding $Lp$

$\supset p$ as a single extra axiom to K has had a long history in modal logic dating from 1937, and is usually referred to simply as T.[7] We shall therefore give the name **T** to the formula itself. In other words, the system T is K +

**T** $Lp \supset p$

This axiom is sometimes called the *Axiom of Necessity*.

All the theorems of K are of course still theorems of T. The derived rules DR1–DR3 and Eq also hold in T. In fact if we look back at how these rules were proved in K, we can see that they are bound to hold in all systems which contain K, provided that they retain the rules US, MP and N. We prove a couple of theorems of T which are not in K.

**T1** $p \supset Mp$

PROOF

| | | |
|---|---|---|
| T[$\sim p/p$] | (1) | $L\sim p \supset \sim p$ |
| (1) × PC | (2) | $p \supset \sim L\sim p$ |
| (2) Def $M$ | (3) | $p \supset Mp$ |

**Q.E.D.**

**T2** $M(p \supset Lp)$

PROOF

| | | |
|---|---|---|
| T1[$Lp/p$] | (1) | $Lp \supset MLp$ |
| K7[$Lp/q$] | (2) | $M(p \supset Lp) \equiv (Lp \supset MLp)$ |
| (1), (2) × Eq | (3) | $M(p \supset Lp)$ |

**Q.E.D.**

We leave it to the reader to show that neither T1 nor T2 is a theorem of K, by defining for each of them a model in which it is false in some world.

The fact that T2 is a theorem of T shows that the following rule, which is a kind of possibility counterpart of N, is not a rule of T:

**P** $\vdash M\alpha \rightarrow\ \vdash \alpha$

The reason is that if P were a rule of T, then from it and T2 we could derive $p \supset Lp$, but as we shall show in a moment, this is not a theorem of T.





### A definition of validity for T

In discussing the modal game on p. 20 we showed that the wff $Lp \supset p$ is valid in every seating arrangement in which all players can see themselves. Transposed into our present frame-theory, this means that **T** is valid on every frame $\langle W,R \rangle$ in which R is *reflexive* – i.e. in which, for every $w \in W$, $wRw$. (We call such frames, for short, *reflexive frames.*) So by theorem 2.2, the system T is sound with respect to the class of all reflexive frames. We shall in fact be able to prove later that T is also complete with respect to this class of frames; so, anticipating this result, we shall say that a wff is *T-valid* iff it is valid on every reflexive frame, and we shall sometimes call a reflexive frame a *T-frame.*

We said a couple of paragraphs back that we would prove that $p \supset Lp$ is not a theorem of T. Now that we have shown that every theorem of T is valid on every reflexive frame, all that we need for this purpose is to find a reflexive frame in which $p \supset Lp$ is not valid. And this is not difficult: imagine a world in which $p$ is true and which can see a world in which $p$ is false, each world being able to see itself.

Since **T** is not K-valid, it is not a theorem of K, and this shows that K and T are distinct systems, with T being a *proper extension* of K.

### The system D

We said on p. 20 that if we interpret $L$ as expressing obligatoriness ('moral necessity') we shall be unlikely to want to regard $Lp \supset p$ as valid, since what it will then mean is that whatever ought to be the case is in fact the case. There is, however, a formula which, like $Lp \supset p$, is not a theorem of K but which with this interpretation it *is* plausible to regard as valid, and that is the wff $Lp \supset Mp$. For if $Lp$ means that it is obligatory that $p$, then $Mp$ will mean that it is permissible that $p$ (not obligatory that not-$p$), and so $Lp \supset Mp$ will mean that whatever is obligatory is at least permissible, which sounds reasonable enough. This interpretation of $L$ is known as a *deontic* interpretation, and for that reason $Lp \supset Mp$ is often called **D**, and the system obtained by adding it to K as an extra axiom is known as the system D;[8] i.e. D is defined as K +

**D** $Lp \supset Mp$

An easily derived theorem of D is

**D1** $M(p \supset p)$

PROOF

| | | | |
|---|---|---|---|
| **PC** | (1) | $p \supset p$ | |
| (1) × N | (2) | $L(p \supset p)$ | |
| **D**$[p \supset p/p]$ | (3) | $L(p \supset p) \supset M(p \supset p)$ | |
| (2), (3) × MP | (4) | $M(p \supset p)$ | Q.E.D. |

In fact D1 would provide an alternative axiom for D, since if we add it alone to K we can derive **D** in the following way:

| | | | |
|---|---|---|---|
| K7$[p/q]$ | (1) | $M(p \supset p) \equiv (Lp \supset Mp)$ | |
| D1, (1) × Eq | (2) | $Lp \supset Mp$ | Q.E.D. |

It is worth noting that if any wff $\alpha$ is a theorem of D, then so is $M\alpha$. For if $\alpha$ is a theorem, N gives $L\alpha$ as a theorem; and then by **D**$[\alpha/p]$ and MP we obtain $M\alpha$.

It is also worth noting that if any system which is an extension of K has any theorems of the form $M\alpha$, that system contains D. To prove this it is clearly sufficient to derive D1 in such a system, and we can do this as follows:

| | | |
|---|---|---|
| Given: | (1) $M\alpha$ | |
| **PC** | (2) $q \supset (p \supset p)$ | |
| (2)$[\alpha/q]$ | (3) $\alpha \supset (p \supset p)$ | |
| (3) × DR3 | (4) $M\alpha \supset M(p \supset p)$ | |
| (1), (4) × MP | (5) $M(p \supset p)$ | Q.E.D. |

In introducing the system D we mentioned that its axiom **D** is not a theorem of K. We shall prove this in a moment, and we shall also prove that **T** is not a theorem of D. **D**, however *is* a theorem of T, since it follows straightforwardly from **T** and T1 by Syll. What this means is that the system D is intermediate between K and T, in the sense that T is a proper extension of D, which in its turn is a proper extension of K.

To find a definition of validity which will match the system D, and also to clarify the difference between D and K, we shall draw attention to a feature of some frames on which we have not so far laid stress. We have observed that not all worlds in a frame need see themselves; but in fact there is nothing in our definition of 'frame' to prevent there being some worlds in a frame which cannot see *any world in that frame at all.* Krister Segerberg has called such worlds *dead ends,*[9] and we shall adopt this terminology in this book. Now the rule [V*L*] says that $L\alpha$ is true in





a world $w$ iff $\alpha$ is true in every world that $w$ can see, and we interpret this to mean that if there is no world at all that $w$ can see, then $L\alpha$ is (trivially) true in $w$, no matter what wff $\alpha$ may be (even if it is $p \wedge \sim p$). (It may be easier to see why we count $L\alpha$ always true in a dead end by seeing why its negation $\sim L\alpha$ is always false in such a world: for $\sim L\alpha$ is equivalent to $M{\sim}\alpha$, and by [V*M*] any wff of the form $M\beta$ can be true in $w$ only if there *is* some world that $w$ can see.) It should now be clear that if a frame contains any dead end $w$, then **D** is not valid on that frame, since in $w$ $Lp$ is true and $Mp$ false, no matter what value is assigned to $p$ there. Since there are such frames, **D** is not K-valid, and is therefore not a theorem of K. A more general consequence is that K has no theorems at all of the form $M\alpha$; for every wff of this form would be invalid on a frame containing any dead end.

Suppose we now consider the class of frames which contain no dead ends, i.e. frames in which each world can see at least one world (itself and/or some other or others). In such frames R is said to be a *serial* relation, and we shall call them *serial frames* for short. In other words, $\langle W,R \rangle$ is a serial frame iff for every $w \in W$, there is some $w' \in W$ such that $wRw'$. Now **D** must be valid on every serial frame: for if it were not, there would have to be a world $w$ in a model based on a serial frame where (i) $Lp$ is true and (ii) $Mp$ is false; but since the frame is serial $w$ must be related to some world $w'$, and then by (i) and [V*L*] $p$ must be true in $w'$ and by (ii) and [V*M*] $p$ must be false there, which is impossible. Since **D** is valid on every serial frame, theorem 2.2 assures us that every theorem of D is valid on every such frame, i.e. that D is sound with respect to the class of all serial frames. We shall be able to prove in Chapter 6 that D is also complete with respect to that class of frames; so, anticipating that result, we now define D-validity by saying that a wff is *D-valid* iff it is valid on every serial frame.

It is now easy to show that **T** is not a theorem of D, and therefore that the system T is a proper extension of D. All we need to do is to exhibit a serial frame on which **T** is not valid, and an example of such a frame is one consisting of two worlds, $w$ and $w'$, where $w$ cannot see itself but can see $w'$, and $w'$ can see itself. **T** is not valid on this frame, for if $p$ is false at $w$ but true at $w'$, then **T** is false at $w$.

### A note on derived rules

Earlier in this chapter we introduced the notation K + Λ to denote the result of adding all the wff in Λ to the basis of K. More generally, where S is any axiomatic modal system containing the transformation rules US, MP and N and $\Lambda$ is any set of wff, we shall let S + $\Lambda$ denote the system obtained by adding all the wff in $\Lambda$ to the basis of S, while retaining the rules US, MP and N. It is a trivial fact that all theorems of S remain theorems of S + $\Lambda$, for the *addition* of new axioms cannot result in the *loss* of any theorems. With derived rules, however, the position is more complicated. We noted earlier on that the rules DR1–DR3 and Eq which we derived in K still hold in all extensions of K. But consider the rule we discussed above and showed not to be a rule of T:

**P** $\vdash M\alpha \rightarrow \; \vdash \alpha$

Now K, as we observed, has no theorems at all of the form $M\alpha$; so P is (trivially) a rule of K. Less trivially, it is also a rule of D. So P is an example of a rule which holds in some systems but not in all their extensions, and this illustrates the care that must be taken with derived rules. If we look back at the way DR1–DR3 and Eq were proved to hold in K, we can easily see why *they* hold in all extensions of K: for they were derived by appealing only to elements in K (theorems and primitive transformation rules) which are still present in all its extensions. But P is a rule of K and D because of features of those systems which are *not* present in all their extensions – in the case of K because the system is too weak to have any theorem satisfying the antecedent of the rule.

So if we are given merely that some rule is a rule of S and that S′ is an extension of S, this does not by itself guarantee that it is also a rule of S′. This is just one of the pitfalls one may encounter in studying axiomatic systems and which should put us on our guard against jumping to conclusions too easily.

### Consistency

We shall say that an axiomatic system is *consistent* iff not every wff is a theorem of that system. In other words, a system is *inconsistent* iff every wff is a theorem. Other definitions of consistency are sometimes given, but provided that the system contains the schema **PC** (or some other way of ensuring that every valid wff of PC is a theorem) and the rules US and MP, all the standard definitions of consistency are equivalent. One such definition is that a system is consistent iff no variable is a theorem. This is equivalent to our definition because (a) if a variable were a theorem, then by US every wff would be one, and (b) if every wff were a theorem, then since $p$ is a wff it would be a theorem. Another definition is that a system is consistent iff no wff and its negation are both theorems. And





this is also equivalent to the definition we have given because (a) if $\alpha$ and $\sim\alpha$ were both theorems, then by substituting $\alpha$ for $p$ and any wff $\beta$ for $q$ in the PC-valid wff $p \supset (\sim p \supset q)$ we could obtain any wff whatsoever as a theorem, and (b) if *every* wff were a theorem, obviously a wff and its negation would both be theorems.

Now clearly the wff $p$ (or any other variable) is not valid on *any* frame; so if a system is sound with respect to any (non-empty) class of frames whatsoever, $p$ is not a theorem of that system, and so the system is consistent. Thus a proof of the soundness of a system is automatically a proof of its consistency.

In Chapter 1 we introduced the notion of an *unsatisfiable* PC wff – i.e. one which has the value 0 for every value-assignment to its variables. It should be obvious that the addition of any unsatisfiable PC wff to any system which contains all valid PC wff and has the rules US and MP would make the system inconsistent; for if $\alpha$ is unsatisfiable, $\sim\alpha$ is valid, and therefore a theorem of the system already, so we should have a wff and its own negation as theorems. But it is also worth noting that if any invalid PC wff at all were a theorem of such a system, the system would be inconsistent. To prove this it will be sufficient, in the light of what we have just said, to show that every invalid PC wff has a substitution-instance which is unsatisfiable, and we can do this as follows:

Let $\alpha$ be any invalid PC wff. The fact that $\alpha$ is invalid means that there is some assignment of truth-values to the variables occurring in it which will give the value 0 to $\alpha$ as a whole. Now let $\alpha'$ be $\alpha$ with $p \vee \sim p$ replacing each variable to which that assignment gives the value 1 and $p \wedge \sim p$ replacing each variable to which it gives the value 0. Then since these two wff have the values 1 and 0 respectively for *every* value-assignment, $\alpha'$ will have the value 0 for every value-assignment – i.e. will be unsatisfiable. But clearly $\alpha'$ is a substitution-instance of $\alpha$.

### Constant wff

In forming $\alpha'$ out of $\alpha$ in the previous paragraph we replaced every variable by a formula whose truth-value could be guaranteed to be 1 or 0 as the case might be, irrespective of any value-assignment made to the variables. A wff of this kind we shall call a *constant* wff. Since the truth-value of $p \wedge \sim p$ does not depend on the truth-value of $p$ ($p \wedge \sim p$ is always false) we may write it as $\perp$ and interpret it as a 'constant false proposition'; and we then define a *constant* wff by saying that $\perp$ is a constant wff, that if $\alpha$ is a constant wff, so are $\sim\alpha$ and $L\alpha$, and that if $\alpha$ and $\beta$ are constant wff, so is $\alpha \vee \beta$. Finally, for convenience, we define the symbol $\top$ as $\sim\bot$, and hence interpret it as a constant true proposition, to be always assigned the value 1.

A constant wff may or may not contain modal operators. A constant PC wff (i.e. one which contains no modal operators but is built up from $\top$ and/or $\bot$ by truth-functional ones only) must have the same truth-value for every value-assignment, and as a result every such wff will be either valid or unsatisfiable. In the case of a constant wff which contains modal operators, its truth-value in any world in a model will not depend on the value-assignment given to variables in that model, but only on how that world is related to other worlds (or to itself) in that model. We shall find further use for constant wff in later chapters.

**Exercises - 2**

**2.1** Prove in K:

(a) $(L(p \supset q) \wedge L(q \supset r)) \supset L(p \supset r)$
(b) $L(p \supset q) \supset (Mp \supset Mq)$
(c) $(L(p \supset q) \wedge M(p \wedge r)) \supset M(q \wedge r)$
(d) $M(p \supset (q \wedge r)) \supset ((Lp \supset Mq) \wedge (Lp \supset Mr))$
(e) $M(p \supset p) \supset (Lq \supset Mq)$
(f) $(Lp \wedge M(q \supset r)) \supset (L(p \supset q) \supset M(p \wedge r))$
(g) $(Lp \wedge Mq) \supset M(p \wedge q)$

**2.2** (a) Let the axiomatic basis of K* be the same as for K except that N is replaced by the axiom $L$T: $L(p \supset p)$, and the rule

R* $\vdash \alpha \supset \beta \rightarrow\ \vdash L\alpha \supset L\beta$ (R* is DR1 but taken as a primitive transformation rule). Show that K and K* have the same theorems.

(b) Let K** be K but with N and **K** replaced by $L$T, R* and K2* $(Lp \wedge Lq) \supset L(p \wedge q)$ (K2* is K2 but taken as an axiom). Show that K and K** have the same theorems.

**2.3** Let T* be the same as T except that in place of **K**, T* contains

K* $L(L(p \supset q) \supset (Lp \supset Lq))$

and in place of N, T* contains R*. Show that T and T* have the same theorems.

**2.4** Prove that K has no theorems of the form $LM\alpha$.

**2.5** Where T′ is exactly like T except that in place of **T** it has

**T′** $p \supset Mp$,

prove that T and T′ have the same theorems.





**2.6** (a) Prove that the following is a rule of K:
$\vdash \alpha \vee \beta \rightarrow\ \vdash M\alpha \vee L\beta$
(b) Prove that the following is a rule of D but not of K:
$\vdash \alpha \vee \beta \rightarrow\ \vdash M\alpha \vee M\beta$
(c) Prove that the following is a rule of T but not D:
$\vdash \alpha \vee \beta \rightarrow\ \vdash M\alpha \vee \beta$

**2.7** Prove in D
(a) $M{\sim}p \vee M{\sim}q \vee M(p \vee q)$
(b) ${\sim}L(Lp \wedge L{\sim}p)$

**2.8** Show that T2 is not a theorem of D.

**2.9** Show that if $M\alpha$ is D-valid then so is $\alpha$.

**2.10** Prove that $\vdash L\alpha \rightarrow\ \vdash \alpha$ is a rule of K and D. [Hint (Chellas 1980, p. 124): For any wff $\alpha$ let $\sigma(\alpha)$ be obtained from $\alpha$ by deleting every modal operator ($L$ or $M$) which is *not* in the scope of another modal operator, and show that any proof of $\alpha$ in K (D) can be converted into a proof of $\sigma(\alpha)$ in the same system.]

**2.11** Let L⁻ be the rule
$\vdash L\alpha \supset L\beta \rightarrow\ \vdash \alpha \supset \beta$
Show that L→ preserves validity in K and D but not in T.

**Notes**

[1] This name, which has now become standard, was given to the system in Lemmon and Scott 1977, p. 29, in honour of Saul Kripke, from whose work the way of defining validity for modal logic which we have begun to describe and will elaborate later is mainly derived. We give the same name to the system and to the formula which is its characteristic axiom, and shall do so for some other systems also. In such cases we shall use bold-face type when referring to the formula, but roman type when referring to the system.

[2] An *axiom* is a specific wff; an *axiom schema* is a statement to the effect that any wff satisfying certain conditions is an axiom. The fact that the axiom schema **PC** gives us infinitely many axioms does not conflict with our requirements for a satisfactory set of axioms, since we have (in the truth-table method, for example), an effective way of determining whether any given wff is a valid wff of PC or not. Although **PC** appeals to a notion of validity it is only PC-validity and makes no reference to the modal operators. It is of course possible to study PC itself as an axiomatic system with a finite number of axioms. See p. 210

below.

[3] The word 'frame' in this sense seems to have been first used in print in Segerberg 1968b, but Segerberg has informed us that the word was suggested to him by Dana Scott. Lemmon and Scott 1977 called frames 'world systems'. Kripke 1963a used the term 'model structure' in a related but not quite identical sense. At this point it might be worth stressing again that the nature of the 'worlds' does not affect the logic. In fact if we take any frame and make an isomorphic 'duplicate', in which the duplicate worlds are related exactly as the originals are, we clearly validate exactly the same formulae.

[4] The symbol $\in$ simply means 'is a member of'. This is a convenient use of set-theoretical notation which we shall employ in this book. Another piece of notation we have been using is the angle brackets $\langle$ and $\rangle$ as in $\langle W,R \rangle$ to indicate the ordered pair of W and R – W and R in that order – or an ordered triple as in $\langle W,R,V \rangle$ and so on. (This contrasts with the use of curly brackets as in $\{a,b\}$ to denote the *unordered* class whose members are precisely $a$ and $b$ without commitment to any order. Thus $\{a,b\}$ is the same class as $\{b,a\}$, $\{a,a\}$ is the same class as $\{a\}$, and so on.) We shall explain other set-theoretical terminology as we proceed.

[5] In calling theorem 2.1 a *theorem* we must be careful not to confuse it with a theorem of K. The theorems of K are the wff which can be derived from the axioms of K by the transformation rules. Theorem 2.1 states a fact *about* K and we prove it by ordinary reasoning. Some authors would call it a *metatheorem* but no confusion ought to arise over the difference in status between theorems like K1–K9 say, and theorems like theorem 2.1.

[6] Where $\Lambda$ is finite K + $\Lambda$ is said to be *finitely axiomatizable*. A system which is not finitely axiomatizable is discussed on p. 185. To call K + $\Lambda$ *axiomatizable* it is often required that $\Lambda$ be effectively specifiable.

[7] Feys 1937 (*vide* esp. pp. 533–535). Feys' own name for the system is 't' (it was first called 'T' by Sobociński 1953). Feys derived the system by dropping one of the axioms in a system devised by Gödel 1933 (p. 39), with whom the idea of axiomatizing modal logic by adding to PC originates. Sobociński (op. cit.) showed that T is equivalent to the system M of von Wright 1951; for this reason 'M' is often used as an alternative name for T. In this book we shall usually refer to systems by names which have become standard, but it might be worth referring, at this point, to an alternative naming system found in Chellas 1980 in the spirit of Lemmon and Scott 1977. This consists in simply listing the axioms in sequence. So T would strictly speaking be KT.

[8] This name is found on p. 50 of Lemmon and Scott 1977.

[9] Segerberg 1971, p. 93.



# 3

# THE SYSTEMS S4, S5, B, TRIV AND VER

In the previous chapter T was the strongest of the systems we discussed. We saw that there are senses of 'necessary' and 'possible' for which some of its theorems seem unacceptable. Nevertheless it seems plausible to hold that there is also a perfectly good and standard sense of these terms in which all the theorems of T are non-controversial and formulae which are not among its theorems – for instance $Lp \supset LLp$ – are at least perplexing.

### Iterated modalities

One feature of $Lp \supset LLp$ and of many other formulae which makes them hard to pronounce on from an intuitive point of view is that they contain consecutive sequences of modal operators; $Lp \supset LLp$, for example, contains the sequence $LL$. Such sequences are known as *iterated modalities*. Now not all formulae containing iterated modalities raise difficulties. If we accept the validity of $Lp \supset p$ (T), for instance, we are not likely to have any qualms about $LLp \supset Lp$ or $LMp \supset Mp$, since they are simply substitution-instances of it. But when we ask, informally, whether $Lp \supset LLp$ is valid, the issue we are raising is this: is whatever is necessary necessarily necessary? when something is necessarily so, is the fact that it is necessarily so always itself something that is necessarily so? Now this is both a disputed question and one of some obscurity, for it is not at all clear under what conditions we should say that something is necessarily necessary. It is, however, at least a reputable and plausible view that in certain well-established senses of 'necessary' it should be answered in the affirmative; it is, for example, plausible to maintain that whenever a proposition is *logically* necessary, this is never a matter of accident but is always something which is logically bound to be the case. We do not, however, need to try to settle the issue definitely here; for what we have just said about $Lp \supset LLp$ is enough to give us a motive for constructing a system stronger than T, in which that formula would be a theorem, and for seeing what such a system would be like.

We have already noted that $LLp \supset Lp$ is a substitution-instance of T, and is therefore a theorem of T and all its extensions; so the new system would have $Lp \equiv LLp$ as a theorem. An equivalential theorem such as this, which entitles us to replace some sequence of modal operators by a shorter sequence, we shall call a *reduction law* of any system of which it is a theorem. Taking the reduction law $Lp \equiv LLp$ as valid would be one way of resolving the perplexity about 'necessarily necessary', for we should then say that $p$ is necessarily necessary whenever $p$ is necessary, and not otherwise. An extension of T such as we are now contemplating would reflect, among other things, the decision to say just this.

Of the various equivalences which could act as reduction laws and have a certain plausibility under many of our intended interpretations of $L$ and $M$, the most important are the following:

**R1** $Mp \equiv LMp$
**R2** $Lp \equiv MLp$
**R3** $Mp \equiv MMp$
**R4** $Lp \equiv LLp$

We shall prove a little later that none of these is a theorem of T; in fact one important feature of T is that it contains no reduction laws whatsoever.[1] If we want to have an extension of T in which R1–R4 are theorems, however, we do not need to go as far as adding them all as new axioms, for three reasons:

1. As we have already mentioned, $LLp \supset Lp$ and $LMp \supset Mp$ are theorems of T itself, and obvious substitutions in T1 will give $Lp \supset MLp$ and $Mp \supset MMp$. So one half of each equivalence is in T already, and it would therefore be sufficient to add the converses, viz.

**R1a** $Mp \supset LMp$
**R2a** $MLp \supset Lp$
**R3a** $MMp \supset Mp$
**R4a** $Lp \supset LLp$





2. Secondly, from R4a we could derive R3a and vice versa, and from R1a we could derive R2a and vice versa. (These derivations are given below.) So it would be sufficient to add as axioms one from each pair, say R1a and R4a.

3. Thirdly, R4a is derivable from R1a, though R1a is not derivable from R4a. (This derivation is also given below.) So we could obtain all four reduction laws by adding R1a to T, while by merely adding R4a we could obtain two of the reduction laws (R3 and R4) but not the other two.

All this suggests the construction of two axiomatic systems, each stronger than T and one of them stronger than the other. The first of these, obtained by adding $Lp \supset LLp$ (R4a) as a new axiom to T, is known as the system S4. The second, obtained by adding $Mp \supset LMp$ (R1a) to T, is known as the system S5.[2]

As in the previous chapter we number theorems using the name of the relevant system; but for theorems of S4 and S5, to avoid confusion, we enclose the theorem number in brackets, writing 'S4(1)' instead of 'S41' and so forth.

## The system S4

The basis of S4 is that of T with the single extra axiom

4 $Lp \supset LLp$

We now prove some theorems.

**S4(1)** $MMp \supset Mp$

PROOF

| | | | |
|---|---|---|---|
| 4[$\sim p/p$] | (1) | $L\sim p \supset LL\sim p$ | |
| (1) × LMI | (2) | $\sim Mp \supset \sim MMp$ | |
| (2) × PC15(Transp) | (3) | $MMp \supset Mp$ | Q.E.D. |

**S4(2)** $Lp \equiv LLp$ [R4]

PROOF

| | | | |
|---|---|---|---|
| T[$Lp/p$] | (1) | $LLp \supset Lp$ | |
| 4, (1) × PC5 | (2) | $Lp \equiv LLp$ | Q.E.D. |

**S4(3)** $Mp \equiv MMp$ [R3]

PROOF

| | | | |
|---|---|---|---|
| T1[$Mp/p$] | (1) | $Mp \supset MMp$ | |
| (1), S4(1) × PC5 | (2) | $Mp \equiv MMp$ | Q.E.D. |

**S4(4)** $MLMp \supset Mp$

PROOF

| | | | |
|---|---|---|---|
| T[$Mp/p$] | (1) | $LMp \supset Mp$ | |
| (1) × DR3 | (2) | $MLMp \supset MMp$ | |
| (2), S4(1) × Syll | (3) | $MLMp \supset Mp$ | Q.E.D. |

**S4(5)** $LMp \supset LMLMp$

PROOF

| | | | |
|---|---|---|---|
| T1[$LMp/p$] | (1) | $LMp \supset MLMp$ | |
| (1) × DR1 | (2) | $LLMp \supset LMLMp$ | |
| (2), S4(2) × Eq | (3) | $LMp \supset LMLMp$ | Q.E.D. |

**S4(6)** $LMp \equiv LMLMp$

PROOF

| | | | |
|---|---|---|---|
| S4(4) × DR1 | (1) | $LMLMp \supset LMp$ | |
| S4(5), (1) × PC5 | (2) | $LMp \equiv LMLMp$ | Q.E.D. |

**S4(7)** $MLp \equiv MLMLp$

PROOF

| | | | |
|---|---|---|---|
| S4(6)[$\sim p/p$] | (1) | $LM{\sim}p \equiv LMLM{\sim}p$ | |
| (1) × LMI | (2) | ${\sim}MLp \equiv {\sim}MLMLp$ | |
| (2) × PC | (3) | $MLp \equiv MLMLp$ | Q.E.D. |

### Modalities in S4

We define a *modality* as any unbroken sequence of zero or more monadic operators ($\sim$, $L$, $M$). We express the zero case by writing '—'. Examples of modalities are: —; $\sim$; $L$; $M\sim$; $LL$; ${\sim}ML{\sim}M$. It is clear, however, that in any system containing LMI every modality can be expressed either without any negation signs at all or else with only one, and that at the





beginning. We shall say that a modality expressed in this way is in *standard form*, and from now on we shall assume that all modalities are expressed in standard form. A modality is said to be an *iterated modality* iff it contains two or more modal operators; thus $LL$ and ${\sim}MLM$ are iterated modalities, but $\sim$ and ${\sim}L$ are not. A modality is *affirmative* if it contains no negation signs and *negative* if it does contain one.

We say that two modalities, $A$ and $B$, are *equivalent* in a given system iff the result of replacing $A$ by $B$ (or $B$ by $A$) in any formula is always equivalent in that system to the original formula; otherwise we say that they are *non-equivalent*, or *distinct* in that system. In a system containing the rules US and Eq the modalities $A$ and $B$ are equivalent iff ($Ap \equiv Bp$) is a theorem of that system. If $A$ and $B$ are equivalent in a certain system, and $A$ contains fewer modal operators than $B$, then $B$ is said to be *reducible* to $A$ in that system. Clearly the formulae we have called reduction laws express the reducibility of certain modalities to others in systems of which they are theorems.

We are now in a position to prove an important result about S4, viz. that in it every modality is equivalent to one or other of the following or their negations:

(i) —; (ii) $L$; (iii) $M$; (iv) $LM$; (v) $ML$; (vi) $LML$; (vii) $MLM$

The proof is straightforward. We ignore the negative cases to begin with. Then clearly (ii) and (iii) are the only one-operator modalities. Now theorems S4(2) and S4(3) entitle us to replace $LL$ by $L$ and $MM$ by $M$; so if we add a modal operator to (ii) or (iii) we shall obtain either a modality equivalent to the original or else (iv) or (v), which are therefore the only irreducible two-operator modalities. In just the same way, if we add a modal operator to (iv) or (v), the only three-operator modalities we can obtain are (vi) and (vii). If, however, we add a modal operator to (vi) or (vii), the result is always equivalent either to the original as before, or else to (iv) or (v) by S4(6) or S4(7); hence there cannot be any irreducible modalities with four or more operators.

Clearly the negative cases can be dealt with in the same way; so what we have shown is that there are at most fourteen distinct modalities in S4. In fact all fourteen are distinct from one another, though we are not yet in a position to prove this.

If we prefix a modality to a wff, $\alpha$, the result is of course itself a wff. The implication relations which hold (in S4) among the formulae thus obtained from (i)–(vii) are set out in the following diagram.[3] (Implication is symbolized by an arrow for typographical convenience.)

$$\begin{tikzcd}
 & & L\alpha \arrow[dl] \arrow[ddrr] & & \\
 & LML\alpha \arrow[dl] \arrow[dr] & & & \\
ML\alpha \arrow[dr] & & LM\alpha \arrow[dl] & & \alpha \arrow[ddll] \\
 & MLM\alpha \arrow[dr] & & & \\
 & & M\alpha & &
\end{tikzcd}$$

We can obtain an analogous diagram for the negative cases by negating all the formulae and reversing the direction of all the arrows.

The situation is strikingly different in T. The absence of any reduction laws in that system means that no matter how many modal operators a modality may contain, we can always construct a longer one which will not be equivalent to it. T therefore contains an infinite number of distinct modalities.

### Validity for S4

We remarked earlier, though without proof, that the S4 axiom **4** ($Lp \supset LLp$) is not a theorem of T. We shall now prove this. We have already shown that every theorem of T is T-valid, i.e. valid on every reflexive frame; so in order to show that $Lp \supset LLp$ is not a theorem of T it is sufficient to describe a reflexive frame on which it is not valid. Here is one such frame: W consists of three worlds $w_1$, $w_2$ and $w_3$. Each world can see itself, $w_1$ can see $w_2$, $w_2$ can see $w_3$, but $w_1$ cannot see $w_3$. Now let $p$ be true in $w_1$ and $w_2$ but false in $w_3$. Then since $w_1$ can see only itself and $w_2$, at both of which $p$ is true, $V(Lp,w_1) = 1$. But since $w_2$ can see $w_3$, at which $p$ is false, $V(Lp,w_2) = 0$. Hence, since $w_1$ can see $w_2$, $V(Lp \supset LLp,w_1) = 0$. So **4** is invalid on at least one reflexive frame, and therefore is not a theorem of T.

A feature of the frame we have just considered which was crucial to falsifying **4** on it was that although in it we had $w_1Rw_2$ and $w_2Rw_3$, we did *not* have $w_1Rw_3$; i.e. the frame was not a transitive one. (A frame $\langle W,R\rangle$





is *transitive* iff R is a transitive relation over W, i.e. iff for any three worlds $w$, $w'$ and $w''$ in W (distinct or identical), if $wRw'$ and $w'Rw''$, then $wRw''$.) And in fact it is impossible to falsify **4** on any transitive frame. The proof is this. Suppose there is some transitive frame $\langle W,R\rangle$ in which for some $w \in W$, $V(Lp \supset LLp,w) = 0$. Then by [V⊃],

(i) $V(Lp,w) = 1$

and

(ii) $V(LLp,w) = 0$.

From (ii), by [V*L*], there is some $w' \in W$ such that $wRw'$ and

(iii) $V(Lp,w') = 0$

and from (iii) in turn there is some $w'' \in W$ such that $w'Rw''$ and

(iv) $V(p,w'') = 0$.

But since R is transitive, we have $wRw''$, and therefore, from (i),

(v) $V(p,w'') = 1$

which contradicts (iv). This proves that **4** is valid on every transitive frame.

*Now the system S4 is K with the two additional axioms* ***T*** *and* ***4***. We showed earlier that **T** is valid on every reflexive frame, and we have now shown that **4** is valid on every transitive frame. So by theorem 2.2 on p. 39, it follows that every theorem of S4 is valid on every frame which is both reflexive and transitive, i.e. that S4 is sound with respect to the class of all such frames. We shall prove in Chapter 6 that S4 is also complete with respect to that class, so we shall define S4-validity as validity on all reflexive and transitive frames, and we shall call any reflexive transitive frame an *S4-frame*. In terms of the modal game, this means that we shall count a wff as S4-valid iff it is valid in every seating arrangement in which whenever any player A can see a player B and B can see a player C, then A must be able to see C.

## The system S5

The basis of S5 is that of T plus the additional axiom

**E** $Mp \supset LMp$

This is the formula we previously called **R1a**.[4] The first three theorems of S5 are proved in the same way as S4(1)–S4(3), but using **E** instead of **4**, and we leave the proofs to the reader. These theorems are

| | | |
|---|---|---|
| **S5(1)** | $MLp \supset Lp$ | |
| **S5(2)** | $Mp \equiv LMp$ | [**R1**] |
| **S5(3)** | $Lp \equiv MLp$ | [**R2**] |

The S4 axiom $Lp \supset LLp$ is not an axiom of S5, but we now prove that it is a theorem of S5. Since the two systems have the rest of their bases in common, this constitutes a proof that S5 contains S4.

**4** $Lp \supset LLp$

PROOF IN S5

| | | | |
|---|---|---|---|
| T1[$Lp/p$] | (1) | $Lp \supset MLp$ | |
| S5(2)[$Lp/p$] | (2) | $MLp \equiv LMLp$ | |
| (1), (2) × Eq | (3) | $Lp \supset LMLp$ | |
| (3), S5(3) × Eq | (4) | $Lp \supset LLp$ | Q.E.D. |

**S5(4)** $L(p \vee Lq) \equiv (Lp \vee Lq)$

PROOF

| | | | |
|---|---|---|---|
| K9[$Lq/q$] | (1) | $L(p \vee Lq) \supset (Lp \vee MLq)$ | |
| (1), **R2** × Eq | (2) | $L(p \vee Lq) \supset (Lp \vee Lq)$ | |
| K4[$Lq/q$] | (3) | $(Lp \vee LLq) \supset L(p \vee Lq)$ | |
| (3), **R4** × Eq | (4) | $(Lp \vee Lq) \supset L(p \vee Lq)$ | |
| (2), (4) × PC5 | (5) | $L(p \vee Lq) \equiv (Lp \vee Lq)$ | Q.E.D. |

**S5(5)** $L(p \vee Mq) \equiv (Lp \vee Mq)$

PROOF

| | | | |
|---|---|---|---|
| S5(4)[$Mq/q$] | (1) | $L(p \vee LMq) \equiv (Lp \vee LMq)$ | |
| (1), **R1** × Eq | (2) | $L(p \vee Mq) \equiv (Lp \vee Mq)$ | Q.E.D. |





**S5(6)** $M(p \wedge Mq) \equiv (Mp \wedge Mq)$

PROOF

| | | | |
|---|---|---|---|
| S5(4)[$\sim p/p, \sim q/q$] | (1) | $L(\sim p \vee L\sim q) \equiv (L\sim p \vee L\sim q)$ | |
| PC | (2) | $(p \equiv q) \supset (\sim p \equiv \sim q)$ | |
| (1) × (2) | (3) | $\sim L(\sim p \vee L\sim q) \equiv \sim(L\sim p \vee L\sim q)$ | |
| (3) × LMI | (4) | $M\sim(\sim p \vee \sim Mq) \equiv \sim(\sim Mp \vee \sim Mq)$ | |
| (4), Def ∧ | (5) | $M(p \wedge Mq) \equiv (Mp \wedge Mq)$ | Q.E.D. |

**S5(7)** $M(p \wedge Lq) \equiv (Mp \wedge Lq)$

PROOF

| | | | |
|---|---|---|---|
| S5(6)[$Lq/q$] | (1) | $M(p \wedge MLq) \equiv (Mp \wedge MLq)$ | |
| (1), **R2** × Eq | (2) | $M(p \wedge Lq) \equiv (Mp \wedge Lq)$ | Q.E.D. |

We can also show that **E** is not a theorem of S4, and therefore that S5 *properly* contains S4. To do this it is sufficient to produce a frame which is reflexive and transitive (and is therefore a frame for S4) on which **E** can be falsified. Such a frame is the frame ⟨W,R⟩ where W consists of two worlds, $w_1$ and $w_2$; each can see itself, $w_1$ can see $w_2$, but $w_2$ cannot see $w_1$. Now let V be a value-assignment which makes $p$ true in $w_1$ but false in $w_2$. Then by [V$M$], since $w_1$ can see itself and $p$ is true there, $V(Mp,w_1) = 1$. But since $w_2$ is the *only* world $w_2$ can see and $p$ is false there, [V$M$] gives us $V(Mp,w_2) = 0$; so by [V$L$], since $w_1$ can see $w_2$, $V(LMp,w_1) = 0$. Thus at $w_1$ $Mp$ is true but $LMp$ is false, and hence $Mp \supset LMp$ is false. So **E** is not a theorem of S4.

## Modalities in S5

We have shown that all the four reduction laws mentioned earlier are theorems of S5. We repeat them here for convenience:

| | | |
|---|---|---|
| **R1** | $Mp \equiv LMp$ | [S5(2)] |
| **R2** | $Lp \equiv MLp$ | [S5(3)] |
| **R3** | $Mp \equiv MMp$ | [S4(3)] |
| **R4** | $Lp \equiv LLp$ | [S4(2)] |

A simple way of summarizing these laws is this: in any pair of adjacent modal operators we may delete the first. Since this procedure may be repeated indefinitely, we have the more comprehensive rule that in any sequence of modal operators we may (in S5) delete all but the last.

It is a straightforward consequence of this that S5 contains at most six distinct modalities, viz.

(i) —; (ii) $L$; (iii) $M$

and their negations. In fact these six modalities are all distinct from one another.

### Validity for S5

If we look back at the frame we used a few paragraphs back to falsify **E**, we can see that although it is reflexive and transitive, it contains a world $w_1$ which can see a world $w_2$, where $w_2$ cannot see $w_1$. This means that the frame is not a *symmetrical* one, since a relation is said to be symmetrical iff whenever it holds in one direction it also holds in the other. I.e., a frame $\langle W,R \rangle$ is symmetrical iff, for any $w$ and $w'$ in W, if $wRw'$ then $w'Rw$.

Now **E** cannot be falsified on any frame which is both transitive and symmetrical. For suppose there is a frame $\langle W,R \rangle$ of this kind on which **E** fails. This means that there is a model $\langle W,R,V \rangle$ based on this frame in which for some $w \in W$,

(i) $V(Mp,w) = 1$

and

(ii) $V(LMp,w) = 0$.

From (i), by [V$M$], there is some $w' \in W$ such that $wRw'$ and

(iii) $V(p,w') = 1$

and from (ii), by [V$L$], there is some $w'' \in W$ such that $wRw''$ and

(iv) $V(Mp,w'') = 0$.

Now since $wRw''$ and R is symmetrical, we have $w''Rw$; and then, since $wRw'$ and R is transitive, we have $w''Rw'$. Hence by (iv) and [V$M$], we have

(v) $V(p,w') = 0$





which contradicts (iii).

Now S5 is K with the two extra axioms **T** and **E**. Since we showed earlier that **T** is valid on every reflexive frame, and have now shown that **E** is valid on every transitive symmetrical frame, theorem 2.2 on p. 39 shows that S5 is sound with respect to the class of all frames which are reflexive, transitive and symmetrical. A relation which is reflexive, transitive and symmetrical is known as an *equivalence* relation. Since we shall be able to prove that S5 is also complete with respect to this class of frames, we define *S5-validity* as validity on every equivalence frame, and an *S5-frame* as a frame of this kind.

An everyday example of an equivalence relation is 'has the same height as', and this can be used to illustrate the fact that when such a relation is defined over a class of objects it divides them into a number (though perhaps only one) of self-contained 'equivalence classes'. Thus if 'has the same height as' is defined over a class of human beings, then for each height that any of them has there will be the 'equivalence class' of all and only those who have that height. Within each such equivalence class everyone will have the relevant relation to everyone, but no one will have that relation to anyone in any other equivalence class. To apply this to frames: if in a frame $\langle W,R \rangle$ R is an equivalence relation, this means that every world will be able to see every world in its own equivalence class but no world in any other equivalence class, and hence that we can equally well think of such a frame, not so much as a single frame but as a collection of separate frames, in each of which every world can see every world. And what this amounts to is that we could equally well, and equivalently, define S5-validity as validity on every frame in which R is a *universal* relation, i.e. one which holds between every pair (distinct or identical) of worlds in that frame.

(In terms of the modal game, what this means is that in order to produce a seating arrangement appropriate for S5, we must either let every player see every player without restriction, or else divide the players into segregated groups, in each of which everyone can see everyone but no one can see anyone outside the group. But if we do the latter, we might as well be playing a number of distinct games simultaneously, in each of which everyone can see everyone.)

In evaluating formulae in models based on frames of this kind, we could replace [V$L$] by the simpler rule

[V$L$S5] $V(L\alpha,w) = 1$ if $V(\alpha,w') = 1$ for every $w' \in W$; otherwise $V(L\alpha,w) = 0$.

However, since this simplification can be undertaken only in the case of S5, we shall for the sake of uniformity stick to [VL] and assume that in S5 frames R is an equivalence relation but not necessarily a universal one.

**The Brouwerian system**

A special interest attaches to the following pair of theorems of S5:

**S5(8)** $p \supset LMp$

PROOF

T1, **E** × Syll

**S5(9)** $MLp \supset p$

PROOF

| | | | |
|---|---|---|---|
| S5(8)[$\sim p/p$] | (1) | $\sim p \supset LM\sim p$ | |
| (1) × LMI | (2) | $\sim p \supset \sim MLp$ | |
| (2) × PC15(Transp) | (3) | $MLp \supset p$ | Q.E.D. |

Neither of these theorems is in S4. Indeed, if we were to add either as an extra axiom to S4 we should obtain a system at least as strong as S5. (In fact we should obtain exactly S5.) In the case of S5(8) we need only to substitute *Mp* for *p* and then apply **R3** to obtain the S5 axiom **E**, and the case of S5(9) is not much more complicated. If, however, we were to add either of them to T instead of to S4 we should obtain not S5 but a system which is weaker than S5 and which neither contains nor is contained in S4. This system has been called the *Brouwerian system*, and S5(8) the *Brouwerian axiom*.[5] We shall use 'B' to refer to the system and '**B**' (in bold face) to refer to the axiom.

The following is a derived rule of B (and also, of course, in view of the way in which it is derived) of S5:

**DR4** $\vdash M\alpha \supset \beta \rightarrow \vdash \alpha \supset L\beta$

PROOF

| | | | |
|---|---|---|---|
| Given: | (1) | $M\alpha \supset \beta$ | |
| (1) × DR1 | (2) | $LM\alpha \supset L\beta$ | |
| **B**[$\alpha/p$] | (3) | $\alpha \supset LM\alpha$ | |
| (3), (2) × Syll | (4) | $\alpha \supset L\beta$ | Q.E.D. |





Yet another way of obtaining S5 would be to add DR4 as a primitive transformation rule to S4, without any new axioms; for then, since $MMp \supset Mp$ (S4(1)) is a theorem of S4, DR4 would immediately give us $Mp \supset LMp$ (i.e. **E**).

**Validity for B**

We show first that **B** is valid on every frame in which R is symmetrical. Let ⟨W,R,V⟩ be any model based on any symmetrical frame. Suppose that for some $w \in W$, $V(p,w) = 1$. Now consider any $w'$ such that $wRw'$. Since R is symmetrical, we also have $w'Rw$; and then, since $V(p,w) = 1$, [VM] gives us $V(Mp,w') = 1$. Since this is so for *every* $w'$ such that $wRw'$, $V(LMp,w) = 1$. Thus whenever *p* is true at any world, so is *LMp*, provided that R is symmetrical; and therefore $p \supset LMp$ is valid on every symmetrical frame.

We already know that **T** is valid on every reflexive frame; so, since the system B is K + **T** + **B**, theorem 2.2 gives us the result that B is sound with respect to the class of all frames which are both reflexive and symmetrical. Such frames we shall call *B-frames*, and we define *B-validity* as validity on every B-frame.

Now we have seen that adding **B** to S4 gives S5, and we have also seen that S4 is weaker than S5; and from this it follows that **B** is not in S4, and hence that S4 does not contain the system B. (In fact the model we used to show that **E** is not in S4 can also easily be used to show that **B** is not in S4.) Furthermore, B does not contain S4 either, since **4** fails on the following reflexive and symmetrical (but non-transitive) frame ⟨W,R⟩: W consists of three worlds, $w_1$, $w_2$ and $w_3$. Each world can see itself, and in addition we have $w_1Rw_2$, $w_2Rw_1$, $w_2Rw_3$ and $w_3Rw_2$. It may help to visualize the frame like this:

$$w_1 \Leftrightarrow w_2 \Leftrightarrow w_3$$

– where the arrows represent the accessibility relation, and it is also assumed that each world is related to itself. If we now form a model on this frame by letting $V(p,w_1) = 1$, $V(p,w_2) = 1$ and $V(p,w_3) = 0$, then **4** fails in this model for just the same reasons as it fails in the model we used on p. 56 to show that **4** is not T-valid (*Lp* is true in $w_1$, but it is false in $w_2$ and so *LLp* is false in $w_1$). The only difference between the two cases is that our present frame is symmetrical as well as reflexive, and we have shown that every theorem of B is valid on every such frame.

So B and S4 are independent systems, in the sense that neither contains the other, and yet each lies between T and S5.

### Some other systems

In later chapters we shall discuss other modal systems; we shall see that there are infinitely many of these, and we shall look at some of the general properties of modal systems. But even with the tools already at our disposal we can see how to define some other systems. For instance, instead of adding **4** to T to obtain S4, we could add it merely to K or to D. The resulting systems are often called K4 and KD4 respectively. If we define *K4-frames* as those which are transitive (whether or not they are reflexive), and *KD4-frames* as those which are both serial and transitive, then the results we have proved so far are sufficient to show that all the theorems of K4 are valid on all K4-frames and all the theorems of KD4 are valid on all KD4-frames. It is not difficult to produce a serial and transitive frame on which **T** fails: the frame we used on p. 45 to prove that **T** is not a theorem of D was in fact such a frame. (It was, of course, not reflexive.) This shows that KD4 does not contain T; *a fortiori*, K4 does not contain it either. We have also shown that **4** is not in T. Thus KD4 and T are independent of each other, and so are K4 and T. Moreover, KD4 is a *proper* extension of K4; for the frame which consists of a single dead end is (trivially) transitive and therefore a frame on which every theorem of K4 is valid; but as we saw on p. 45, **D** is not valid on any frame which contains a dead end.

We can similarly add **B** to K or to D instead of to T, to obtain the systems KB and KDB, which can easily be shown to be sound with respect to the classes of symmetrical frames and serial and symmetrical frames respectively. It can then be shown, by arguments of the kind used in the previous paragraph, that each of KB and KDB is independent of each of K4, KD4 and T; but we leave this task to the reader.[6]

### Collapsing into PC

We shall now look at a system which can be obtained by adding even to D, and *a fortiori* to any of the stronger systems we have mentioned, the extra axiom $p \supset Lp$. This formula is not even S5-valid, since it can easily be falsified on a two-world frame in which each world can see both worlds (and which is therefore an S5-frame), by letting $p$ be true in one world but false in the other. Nevertheless adding it even to S5 would not result in an inconsistent system, for the following reason. Consider a frame in which there is only one world, $w$, and it is related to itself. This is clearly an S5-frame, but $p \supset Lp$ is valid on it; for if $V(p,w) = 1$, then





$V(p,w') = 1$ for every $w'$ such that $wRw'$, since there is only one such $w'$, namely $w$ itself, and so $V(Lp,w) = 1$. Every theorem of S5 + $p \supset Lp$ is therefore valid on this frame; but $p$ is not, since there is obviously a value-assignment which makes $p$ false at $w$. So not every wff is a theorem of S5 + $p \supset Lp$ ; i.e. the system is consistent.

In this system the new axiom, together with **D**, immediately yields $p \supset Mp$ (by Syll); from this (by $[\sim p/p]$, Transp and LMI) we can obtain $Lp \supset p$, and then, by simple steps, $Lp \equiv p$ and $Mp \equiv p$. By the rules US and Eq every formula would be then equivalent to the result of deleting all its modal operators; so in any formula we could delete or insert $L$s and $M$s to our heart's content (provided we preserved well-formedness), and the result would be equivalent to the original. In such a system, therefore, the modal operators would merely 'idle'; in interpreting the system we could draw no significant distinction between necessity, possibility and truth, and for all practical purposes it could be regarded simply as the Propositional Calculus itself, encrusted with $L$s and $M$s as mere typographical embellishments. The PC wff which results from deleting all the modal operators in a modal wff $\alpha$ is said to be the *PC-transform* of $\alpha$. A system such as the one we have just described, in which every wff is equivalent to its own PC-transform, may be said to *collapse into PC*.

It is worth noting that although in the previous paragraph we appealed to the rule Eq, we could have obtained all our results from the new axiom and **D** alone (together with the axiom schema **PC**). We did not even need to have **K** as an axiom, nor did we need the rule of Necessitation. Moreover, the system would clearly contain S5, since the results of deleting all the modal operators in **T** and in **E** are PC theorems.

If we add the stronger axiom $p \equiv Lp$ even to K the resulting system similarly collapses into PC. The system D + $p \supset Lp$ (or K + $p \equiv Lp$) is known as the *Trivial* system (*Triv* for short), because in it the modal operators are trivial in the sense we explained earlier. The wff $p \equiv Lp$ is itself sometimes called **Triv**.

It is only in the very strong system Triv that every wff is equivalent to its PC-transform. Even in the much weaker system D, however (and in all systems containing it), there is a somewhat analogous relation between a certain class of wff and their PC-transforms. These are the wff which at the end of the previous chapter we called constant wff - wff constructed out of the constant true and false propositions $\top$ and $\bot$ by truth-functional and modal operators. The PC-transform of any constant wff is of course itself a constant PC wff, and we noted on p. 48 that every such wff is either PC-valid or PC-unsatisfiable. The relation is this: If $\alpha$ is any constant wff, then if its PC-transform is PC-valid, $\alpha$ itself is a theorem of D; otherwise (i.e. if its PC-transform is unsatisfiable) $\sim\alpha$ is a theorem of D. Let us denote the PC-transform of any wff $\alpha$ by $\tau(\alpha)$; then we can state the result as the following lemma:

LEMMA 3.1 Let $\alpha$ be any constant wff. Then if $\tau(\alpha)$ is PC-valid, $\vdash_D \alpha$; otherwise $\vdash_D \sim\alpha$.

Since every constant wff can be constructed from $\bot$ by $\sim$, $\vee$ and $L$, in order to prove the lemma it is sufficient to show (i) that it holds for $\bot$, (ii) that if it holds for a wff $\alpha$ it also holds for $\sim\alpha$, (iii) that if it holds for $\alpha$ it also holds for $L\alpha$, and (iv) that if it holds for $\alpha$ and for $\beta$ it also holds for $\alpha \vee \beta$.

To show (i) we need only remark that the PC-transform of $\bot$ is $\bot$ itself, and that since $\bot$ is unsatisfiable its negation, $\sim\bot$, is PC-valid and therefore a theorem of D. (ii) and (iv) hold by purely PC principles, and we omit the details of their proofs here. (They rely on the fact mentioned above that if $\alpha$ is a constant wff then $\tau(\alpha)$ is either PC-valid or PC-unsatisfiable.) For (iii) the proof is this: (A) Suppose that $\tau(L\alpha)$ is PC-valid. Clearly $\tau(L\alpha)$ is the same wff as $\tau(\alpha)$; so, since the lemma is assumed to hold for $\alpha$, $\vdash_D \alpha$; hence by N, $\vdash_D L\alpha$. (B) Suppose that $\tau(L\alpha)$ is not PC-valid. As before, $\tau(L\alpha)$ is the same wff as $\tau(\alpha)$, and the lemma is assumed to hold for $\alpha$. Hence $\vdash_D \sim\alpha$; hence (by N) $\vdash_D L\sim\alpha$; hence (by **D**) $\vdash_D M\sim\alpha$; hence (by LMI) $\vdash_D \sim L\alpha$.

We shall have a use for lemma 3.1 shortly. In the meantime, however, we shall consider another way in which a system can collapse into PC.

We produced the system Triv by adding $p \supset Lp$ to D. Adding it to K would not have been enough. For consider the frame which consists of a single dead end. It is easy to check that on this frame $p \supset Lp$ is valid but $Lp \supset p$ is not, so the latter is not a theorem of K + $p \supset Lp$. In that system, therefore, unlike Triv, $p \supset Lp$ and $Lp \supset p$ are not equivalent, even though they have the same PC-transform.

The system we are about to consider, however, is not K + $p \supset Lp$ but the even stronger system produced by adding the axiom $Lp$ to K. From this axiom we can of course obtain by US every wff of the form $L\alpha$ as a theorem – even $L\bot$. This system is known as the *Verum* system (*Ver* for short). It no doubt appears bizarre in many ways, and certainly seems to impose some strain on the attempt to interpret $L$ as meaning





'necessarily'. It is nevertheless a consistent system because $Lp$, and therefore every theorem of the system, is valid on the one-world dead end frame we have just referred to, but $p$ is not. In Ver any wff will be equivalent not, as in Triv, to its own PC-transform, but to the wff which results from replacing every well-formed expression of the form $L\alpha$ in it by $\top$, and every one of the form $M\alpha$ by $\bot$. Since the formula thus obtained will always be a PC wff, we could regard the Verum system as providing a different form of collapsing into PC.

The reason for calling K + $Lp$ the *Verum* system is that in interpreting it we think of $L\alpha$ as always true. The wff $Lp$ is sometimes itself called **Ver**.

Triv and Ver are incompatible systems; i.e. the system K + **Triv** + **Ver** is inconsistent. For if both $Lp$ and $p \equiv Lp$ are theorems, so is $p$, and therefore by US every wff is a theorem. Hence Triv is not contained in Ver, nor is Ver in Triv.

Two other results which can be proved about these two systems are:

(1) Every normal modal system, in the sense explained on p. 40 (i.e. every consistent extension of K which retains the rules US, MP and N), is contained either in Triv or in Ver. (Some systems, of course, like K itself or K4 or KB, are contained in both.)

(2) Each of Triv and Ver is a maximal system, in the sense that in the case of each of them, if *any* wff which is not already a theorem were added to it, the resulting system would be inconsistent.[7]

The second of these results follows from the first. To show this, let us suppose that we have proved (1). Then to prove that Triv is a maximal system we take any wff $\alpha$ which is not a theorem of Triv. In that case, the system Triv + $\alpha$ is not contained in Triv, and so by (1) it must either be inconsistent or else be contained in Ver; but the latter would mean that Triv itself is contained in Ver, and we saw above that it is not. That Ver is also a maximal system follows from (1) in an exactly analogous way. So in order to prove both (1) and (2) it will be sufficient to prove (1).

Our strategy for proving (1) will be to prove the following two lemmas, from which (1) clearly follows immediately:

LEMMA 3.2 Every consistent extension of K which is not contained in Ver contains D.

LEMMA 3.3 Every consistent system which contains D is contained in Triv.

The proof of lemma 3.2 will be made easier by some techniques we shall introduce on p. 108, so we shall postpone it till then. Lemma 3.3, however, can be proved with our presently available resources, as follows:

*Proof of lemma 3.3*: It is sufficient to show that if S is any system which contains D and has some theorem $\alpha$ which is not a theorem of Triv, then S is inconsistent. We show this as follows. Since $\alpha$ is not a theorem of Triv, its PC-transform $\tau(\alpha)$ is not PC-valid. Now precisely the same procedure which we used on p. 47 to show that every invalid wff of PC has a substitution-instance which is an unsatisfiable constant wff will also produce for any wff with an invalid PC-transform a substitution-instance which is a constant proposition whose PC-transform is an unsatisfiable wff. Let $\alpha'$ be such a substitution-instance of $\alpha$. Then (1) by US, $\alpha'$ is a theorem of S. But by lemma 3.1, $\sim\alpha'$ is a theorem of D, and hence, since S contains D, it is also a theorem of S. Thus both $\alpha'$ and $\sim\alpha'$ are theorems of S, and S is therefore inconsistent.

**Exercises — 3**

**3.1** Prove in S4:
(a) $L(p \supset q) \supset L(Lp \supset Lq)$
(b) $(Lp \vee Lq) \equiv L(Lp \vee Lq)$
(c) $ML(p \supset LMp)$
(d) $M(Lp \supset Mq) \supset M(p \supset q)$

**3.2** Where A is any affirmative modality (i.e. a string of *L*s and *M*s) show that $L(p \supset q) \supset L(\mathrm{A}p \supset \mathrm{A}q)$ is a theorem of S4.

**3.3** Show that T with $L(p \supset q) \supset L(Lp \supset Lq)$ in place of **K** is deductively equivalent to S4.

**3.4** Prove that the modalities listed on p. 55 are non-equivalent in S4.

**3.5** Prove that where $L^n p$ is $p$ with n *L*s in front of it then for n $\neq$ m, $L^n p \equiv L^m p$ is not a theorem of T.

**3.6** S4.2 is S4 + the axiom
**G1** $MLp \supset LMp$
Prove that S4.2 has only four proper (i.e. non-empty) affirmative modalities, *L*, *ML*, *LM*, and *M* and that in terms of strength they can be





linearly ordered in the order listed here.

**3.7** Prove the following in S5:
(a) $L(Lp \supset Lq) \vee L(Lq \supset Lp)$
(b) $L(Mp \supset q) \equiv L(p \supset Lq)$
(c) $MLp \supset (Mq \supset L(p \wedge Mq))$

**3.8** Show that S5 can be axiomatized as
(a) D + **B** + **E**
(b) S4 + **B**

or as
(c) K +
**E1** $LMLp \supset p$
**E2** $MLp \supset LMLLp$
(d) Show that neither K + $\mathrm{E_1}$ nor K + $\mathrm{E_2}$ on its own gives T, KB (K + $p \supset LMp$) or K4 (K + $Lp \supset LLp$). (Hughes 1980).

**3.9** Show that S5 can be axiomatized as **PC**, US, MP, **T** and $\vdash \alpha \supset \beta \rightarrow \alpha \supset L\beta$, provided $\alpha$ is fully modalized, i.e. every variable in $\alpha$ is in the scope of a modal operator. (Prior 1955a, Lemmon 1956).

**3.10** Show that adding $L(p \vee Lq) \supset (Lp \vee Lq)$ to T gives a system deductively equivalent to S5.

**3.11** Prove that K + **E** is sound with respect to the class of frames in which if $w\mathrm{R}w'$ and $w\mathrm{R}w''$ then $w'\mathrm{R}w''$.

**3.12** Prove in B
(a) $(MLp \wedge MLq) \supset LM(p \wedge q)$
(b) $MLp \supset LMp$

**3.13** Show that B can be axiomatized by dropping N and **K** and adding **B** and
R* $\vdash \alpha \supset \beta \rightarrow \vdash L\alpha \supset L\beta$ (Jennings 1981)

**3.14** Show that $\vdash L\alpha \rightarrow \vdash \alpha$ is not a rule of KB (i.e. K + $p \supset MLp$).

**3.15** Show that if **K** is strengthened to an equivalence $(L(p \supset q) \equiv (Lp \supset Lq))$ then T would collapse into PC.

**3.16** Prove that the addition to S5 of the axiom $LMp \supset MLp$ would make the resulting system collapse into PC.

**3.17** Show that K + $p \supset Lp$ is sound with respect to any class consisting of just two frames, each containing just one world. In one frame this world can see itself. In the other it is a dead end.

**3.18** Set out fully the inductive steps for cases (ii) and (iii) in the proof of lemma 3.1 on p. 66.

**Notes**

[1] See Bellissima 1989. Thomas 1964 cites as an unpublished result by Sobociński the fact that for each n the system $S4_n$, obtained by adding $L^np \supset L^{n+1}p$ (where $L^np$ is $p$ preceded by n $L$s) properly contains $S4_m$ when n < m. The result is easy to obtain using the obvious definition of validity for these systems (Exercise 3.5). Sugihara 1962 proves that **T** + $LLp \supset LLLp$ contains infinitely many distinct modalities.

[2] The names 'S4' and 'S5', which have now for long been standard, derive from Lewis and Langford 1932 (p. 501), where systems deductively equivalent to these are the fourth and fifth in a series of modal systems. (For more on this see Chapter 11.) In the naming system referred to in note 7 on p. 50 S4 would be KT4, S5 would be KTE, and so on.

[3] This diagram is given in Prior 1957, p. 124. The results were originally obtained by Becker 1930 and Parry 1939.

[4] The name **E** for this wff is found on p. 50 of Lemmon and Scott 1977. It corresponds with a condition they call the *euclidian* condition. (See exercise 3.11.) Chellas 1980, p. 6 calls it 5, and thus refers to S5 as KT5.

[5] This formula derives from Becker 1930, p. 509. An alternative version of **B** is of course $\sim p \supset L{\sim}Lp$. Some authors have called **B** the *Brouwersche* axiom, and the system the *Brouwersche* system, perhaps because in Lewis and Langford 1932, p. 497, Becker's German phrase 'Brouwersche Axiom' is quoted untranslated. The name derives from L.E.J. Brouwer, the founder of the intuitionist school of mathematics. In the intuitionist propositional calculus the law of double negation is not valid as an equivalence. More precisely, $p \supset \sim\sim p$ is valid but $\sim\sim p \supset p$ is not. One way of making this sound reasonable has been to suppose that in this calculus $\sim$ means something like 'it is not possible that', i.e. that it means what we usually mean by $L\sim$. Now if we replace $\sim$ by $L\sim$ then $\sim\sim p \supset p$ becomes $L{\sim}L{\sim}p \supset p$, i.e. $LMp \supset p$, and $p \supset \sim\sim p$ becomes $p \supset LMp$, i.e. **B**. On this view **B** therefore represents the intuitionistically





acceptable direction of the double negation law, and so has a connection, albeit somewhat tenuous, with Brouwer. (For a discussion of the intuitionistic propositional calculus see pp. 224–225.)

[6] These systems have some interesting properties in the matter of derived rules. We mentioned on p. 45 that a derived rule may hold in a system but not always in a stronger system. Some interesting examples of this are provided on p. 181f. of Chellas 1980. Thus the rule $\vdash L\alpha \rightarrow \ \vdash \alpha$ is a rule of K, D and KDB but not a rule of KB. (It is trivially a rule of every extension of T). Another example is what is called the 'rule of disjunction' that if $\vdash L\alpha \lor L\beta$ then either $\vdash \alpha$ or $\vdash \beta$. This is a rule of K, D, T and S4 but not a rule of B or S5. (See Chellas 1980, p. 181 and Hughes and Cresswell 1984, pp. 96–100. A weaker version of DR4 on p. 62, viz. $\vdash M\alpha \supset \alpha \rightarrow \ \vdash \alpha \supset L\alpha$, is studied in Chellas and Segerberg 1994 and Williamson 1994. Other studies of the effects of rules in systems of modal logic may be found in Williamson 1988 and 1992, where various philosophical interpretations of $L$ are argued to fit certain rules.

[7] These results are obtained algebraically in Makinson 1971. See also Segerberg 1972. For some early results of this kind see McKinsey 1944.