Possible Models Diagrams — a New Approach to Teaching Propositional Logic

by

Matthew Charles Clarke

Submitted in fulfillment of the academic requirements
for the degree of Master of Science
in the Department of Computer Science and Information Systems

University of Natal

Pietermaritzburg, 1994

Abstract

Possible Models Diagrams (PMDs) are simple graphs which may be used to represent propositional expressions. Technically, these graphs are hypercubes in which the vertices are partitioned into two sets: one representing the possible models in which the propositional expression turns out to be true, and the other representing the possible models in which the propositional expression turns out to be false. PMDs can be used to define boolean operators, to analyse whether a propositional expression is tautological, contingent or inconsistent, and to determine the validity of propositional sequents.

This dissertation describes both theoretical and pedagogical aspects of PMDs, and places this new approach within the context of a course on logic for first year computer science students. In this course, various forms of logical representation are taught and the ability to translate between those representations is emphasised. An extensive comparison is made between PMDs and other methods of teaching propositional logic. In particular, qualitative and quantitative evidence is given to show that students perform better when using PMDs than they do when using truth tables. The advantage of PMDs for the purpose of teaching is that they are iconic: that is, they are symbolic pictures which combine the expressive power of symbolism with the memorability of visual images.

Preface

The work described in this dissertation was carried out in the Department of Computer Science and Information Systems, University of Natal, Pietermaritzburg, between July 1992 and June 1994, under the supervision of Mr Robert Dempster and Dr Diane Grayson.

This study represents original work by the author and has not been submitted in any form for any degree or diploma to any university. Where use has been made of the work of others, it is duly acknowledged in the text.

Matthew C. Clarke

Acknowledgments

My most heartfelt thanks go

to my two supervisors Rob and Diane,

to my wonderful wife Rosalee,

and to the many students who have acted as my guinea-pigs.

Mr J. Levin provided valuable advice on the use of statistical techniques in Chapter 7.

Dr J. Raftery's advice lead to a cleaner distinction between the syntactic and semantic aspects of propositional logic presented in Chapter 2.

Contents

Figures

Tables

Symbols and Abbreviations

Although each of the abbreviations and symbols in the following table are defined in the text, they are all collected here for ease of reference.

Symbol	Meaning
PMD	Possible Models Diagram
wff	Well-Formed Formula
FOPL	First-Order Predicate Logic
$\Rightarrow$	Material Implication
v	Disjunction
&	Conjunction
~	Negation
$\therefore$	Entailment
$\exists$	Existential Quantifier ("there exists")
$\forall$	Universal Quantifier ("for all")
$=_{def}$	Is defined to be
$\cong$	Is isomorphic to (with respect to graphs)

Chapter 1: Introduction

1.1 Historical Context

Most university courses in Computer Science include some component introducing the student to formal logic, and the University of Natal in Pietermaritzburg has been no exception. One quarter of the first year Computer Science course was devoted to "Discrete Structures" — a module which included set theory, propositional and predicate logic as well as an introduction to grammars and finite automata. In recent years it has been noticed that students found this module particularly difficult.

At a departmental meeting during 1992, it was decided to reduce the scope of the Discrete Structures module by dropping the section on grammars and automata, thus allowing an extension of the logic component. The motivation for this change was the department's fourfold belief that:

- A firm foundation in logical thinking is the basis of many concepts and skills in computer science. Not only is formal logic central to the operation of a computer (and hence to understanding the operation of computers), but being able to reason logically should enable students to cope more easily with other areas of study in their computer science degree.
- Changing social patterns in South Africa has lead to an increasing number of educationally disadvantaged students enrolling for university degrees. This heightens the need to explicitly address the process of logical reasoning, rather than hoping that students will either commence their studies with the necessary skills or magically pick them up by osmosis from existing courses.
- An extended module on logic was seen as a good foundation for any academic pursuit and the Department saw this as a contribution to the broader task of educating critical thinkers in the Science Faculty.
- Students would be more intellectually mature when facing the section on grammars and automata in second year. In addition, that section was not seen as essential for those students who were not majoring in computer science.

A detailed description of the new module is given in the Appendix and discussed in Chapter 6. This module is taught over one semester through 26 lecture periods and 13 tutorial periods. The course covers topics common to most first courses on logic in computer science degrees, and also includes a sizable section (about 17% of the course) on informal and inductive logics. Roughly half the course addresses the standard topics of propositional and predicate logics and then a final section provides a brief exposure to other forms of logic such as fuzzy logic and modal logic.

In the process of designing this module, the author developed a novel method of teaching propositional logic which uses simple graphs. These graphs, known as "Possible Models Diagrams" (PMDs), may be viewed as an alternative to Truth Tables. This approach has been used and refined over the past three years and this dissertation is written as a retrospective analysis of the effectiveness of the PMD approach. The dissertation describes the use of PMDs and draws together a variety of qualitative and quantitative data to show that the approach is an effective method of teaching propositional logic to computer science students. The idea behind these diagrams was inspired by a talk given by Prof. J. Heidema in 1992, although the context in which he used them was much different than the didactical application described here (see [BRIN87]).

This research is inter-disciplinary, with roots in both logic and in education. It is hoped that the level of detail on the logical issues is sufficient for readers who come from a background in education, and that the level of detail on educational issues is sufficient for readers who come from a background in logic.

1.2 Objectives

The objectives of this dissertation are thus:

1. To describe the use of PMDs and to show the theoretical soundness of this method of representing propositional expressions;

2. To appraise the pedagogical soundness of PMDs by comparing students' competence in propositional logic when using Possible Models Diagrams with their competence when using truth tables.

3. To place the teaching of PMDs in the context of a course on logic for computer science students.

4. To compare the PMD approach with previous methods of teaching propositional logic in order to situate this approach within a historical context.

1.3 Outline of Dissertation

Chapters 2 and 3 form an extended literature survey covering the range of approaches to the teaching of logic and the types of difficulties which are encountered when teaching logic. Chapter 2 sketches the historical development of logic from Aristotle to First-Order Predicate Logic (FOPL) and the accompanying development of techniques for teaching logic. The central focus is on the variety of proof techniques which have been developed over the past century and a detailed comparison is made of twelve different systems. The later introduction of PMDs can then be seen in the context of these other approaches.

Chapter 3 discusses some typical mistakes made by neophyte logicians which indicate common misconceptions and difficulties with learning logic. These observations provide motivation for improving our teaching methods, and also indicate some directions which such improvements may take. In particular, the problems arising from a truth functional definition of material implication are described and a large number of suggestions for avoiding or overcoming these problems are compared. In the context of the problems with material implication, an approach based on set theory is foreshadowed.

PMDs are first introduced in Chapter 4. This chapter grounds PMDs in graph theory and boolean algebra. Any propositional expression can be represented as a boolean function, and any boolean function can be represented as an induced subgraph of a hypercube. A PMD is just such a graph, drawn in an appropriate form. Algorithms for constructing these induced subgraphs are given as well as a method for using PMDs to ascertain the validity of propositional sequents. PMDs are shown to be informationally equivalent to truth tables. This chapter is a slightly modified version of a paper to appear in the journal of Mathematical and Computer Modeling.

Chapter 5 re-presents some of the material of Chapter 4, but in a friendlier and less theoretical style. This chapter shows the way in which PMDs are introduced to students: instead of the graph theory foundation of the previous chapter, the concepts required for PMDs are drawn from set theory. The logical connectives of propositional calculus are defined and the procedures for building and using PMDs are described. The method of proving propositional sequents, and the equivalence of PMDs and truth tables are again covered, but this time from a didactical rather than theoretical perspective. This chapter is a slightly modified version of a paper presented at the ACM's 24th SIGCSE Technical Symposium on Computer Science Education.

PMDs provide just one more tool for teaching logic, and it is important to understand how such a tool can be effectively integrated into the whole tool kit. Chapter 6 seeks to do this by describing the role of PMDs in a logic course for first year computer science students. The course content is outlined and teaching methods and underlying philosophy are described. This course presents students with a variety of types of logic (inductive, propositional, predicate, modal, fuzzy, multi-valued and probabilistic) and a variety of representational tools (eg English, sets, truth tables, PMDs, and a natural deduction system). The importance of using multiple representations is justified and the significance of being able to translate between representations is emphasised.

In Chapter 7, student responses to this logic course (in particular to PMDs) are analysed. The reader will be aware by now that this research has been theoretical and qualitative rather than quantitative. The PMD approach arose in the midst of teaching logic in a computer science department and as yet no controlled experiments have been conducted to determine the pedagogical effectiveness of PMDs. Nevertheless, data has been recorded over the past three years in the form of student answers to class tests, assignments and exams. This data is summarised in Chapter 7 and statistical analysis has been applied to the extent that it is meaningful. The analysis seeks to establish whether students prefer to use PMDs or truth tables, and also compares their ability to solve problems using PMDs and truth tables. This chapter also contains a selection of actual student responses which provide good examples for qualitative analysis. This qualitative data provides a rich source of insight into student difficulties and misconceptions, and has been invaluable for the development and improvement of teaching methods.

In addition to the data extracted from assessment of students, Chapter 7 also discusses the students' assessment of the logic course, based on responses to course evaluation questionnaires. These responses paint a general picture of students who are satisfied with both the knowledge they have acquired and the means by which that knowledge was imparted.

Overall conclusions are brought together in Chapter 8 in two sections: technical conclusions regarding the theoretical soundness of PMDs as a tool for propositional logic; and pedagogical conclusions regarding the effectiveness of PMDs as a tool for teaching propositional logic.

A lengthy Appendix gives a more detailed description of the complete logic course which was outlined in Chapter 6. This Appendix is a copy of the notes which are distributed to students at the beginning of the course.

1.4 Citation Style

In parts of this dissertation I have tried to indicate the historical flow of ideas in logic and have chosen a style of citation to emphasise this flow. The teaching of logic has changed drastically over the last hundred years and I want to present the PMD approach in the light of these changes. The citation style used throughout this dissertation was chosen because it makes it convenient to indicate original publication dates. Thus, a citation such as [HILB28 p125] refers to work by Hilbert and Ackermann which was *first published* in German in 1928, though this may not have been the edition available to me. The Bibliography indicates that the actual edition being referenced by [HILB28] is the 1950 English translation and the cited page number ("p125" in this example) refers to that edition rather than the original German edition. The year of publication indicated in the citation can be assumed to be in the 20th century, unless otherwise indicated.

In addition to the normal citation of books and journals, the dissertation also cites various electronic sources. In particular I have received assistance from the Internet sci.logic newsgroup and the logic-l distribution list[1] (an international group of people interested in the education of logic). I cite such sources in footnotes and where possible provide the source's e-mail address.

---

1 One can become part of this interest group by sending the message "SUBSCRIBE LOGIC-L" to the address <listserv@bucknell.edu>.

Chapter 2: Historical Background to Logical Methods

2.1 Historical Trends in the Teaching of Logic

Between Aristotle and the late nineteenth century, formal logic meant categorical syllogisms. This is reflected in textbooks on elementary logic up until the rise of mathematical logic which started to find its way into textbooks after 1920.

Around the turn of the century books such as [SIGW1873, KEYN1884, FOWL1895, WELT1896, READ1898, CREI1898, MELL04, RUSS06, JOSE06, WELT11, RUSS14, and WELT20] promoted a common syllabus covering the basic laws of thought, definitions and classifications, types of propositions (categorical, hypothetical and disjunctive), lists of common logical fallacies, and the standard syllogistic forms. Quite a few make use of either Euler's Circles or Venn Diagrams (see Section 2.3.6) to help explain categorical statements and valid forms of categorical inference. Most also have a section on inductive forms of reasoning, including generalisation, causation, the role of observation and hypotheses in science. Some also discuss the use of probability and statistics.

The use of Venn Diagrams as an aid to understanding valid forms of inference is noteworthy, since the approach taken by this thesis revives that technique in part. The earlier Euler Circles provided different diagrams for each of the four forms of categorical proposition, but Venn proposed that a single diagram could be used to represent all four and more [VENN1881]. Whereas Venn Diagrams have been used in recent primary and secondary school syllabi as a way of describing set theory, their original purpose was to represent categorical relationships. In the following chapters it will be seen that Venn Diagrams serve an excellent role in linking logic to set theory so that students who know set theory can more readily learn the concepts of formal logic.

Through the second half of the nineteenth century and early decades of the twentieth century many advances were being made by Peirce [HART74], Shröder (and even earlier by Boole [BOOL1854]), Peano, Frege, Russell and Whitehead [WHIT10], Hilbert [HILB28] and others, but these were not immediately incorporated into standard courses on logic. The change from Classical logic to Modern mathematical logic took some time and even as late as 1956 some authors resisted the change[1]. Nevertheless, the importance of syllogisms gradually gave way to the more comprehensive mathematical logic.

---

[1] I refer here to a remarkable book whose third edition in 1956 still refused to drop the Classical approach. The Jesuit author sought to defend Scholasticism against the scourge of modern heresy [JOYC56].

Many books published during this period gave equal attention to the Classical and the Modern [STEB30, EATO31, STEB43, WERK48] but in 1928 Hilbert and Ackermann published "Principles of Mathematical Logic" [HILB28], a German textbook in which Classical Logic took a back-seat. They described a sentential calculus (what we now call "propositional calculus") and a functional calculus (which later editions renamed "predicate calculus"). Their approach was axiomatic, allowing the truth of an expression to be established by rewriting the expression in a normal form, rather than employing truth tables. Aristotelian logic was shown to be a subset of this new mathematical logic.

From 1930 to 1950 the topics included in logic textbooks (and by implication in university logic courses) varied considerably, but over these decades a new syllabus was gradually assembled to replace the Aristotelian syllabus. This new syllabus largely ignores inductive reasoning and instead focuses on the elements of deductive inference —

- Definitions of logical connectives for negation, conjunction, disjunction and material implication (usually by truth table);
- Definition of well-formed formulae (wffs) based on those connectives;
- The conversion of natural language statements into symbolic form;
- The use of truth tables to classify wffs as either tautological, contingent or inconsistent;
- Some method (see Section 2.3) whereby the validity or derivability of propositional sequents may be established;
- Proof, or at least some discussion, of the consistency and completeness of the propositional calculus;
- The syntax and semantics of predicates and quantifiers;
- Extending the sequent validation technique to include predicates and quantifiers;
- The role of interpretations in the validation of predicate sequents; and
- Proof, or at least some discussion, of the consistency and completeness of the predicate calculus.

The topics included in this list constitute what is called First Order Predicate Logic (FOPL). This syllabus has continued to be fairly standard up to the present time (see for example the recommendations of the Association of Symbolic Logic in [ASL94]), though other topics such as identity, modal logic, and normal forms are often also included in a first course in logic at university level. There are, however, two points on which there is considerable divergence of opinion: which symbols should be used for the logical connectives, and which method should be used for proving or deriving sequents.

2.2 Choice of Symbolism

Virtually all authors now use an infix notation rather than the earlier prefix notation of Lukasiewicz[2]. However, the choice of which symbols to use has not been standardised. Table 2.1 shows the range of symbols commonly used. The first symbol listed for each operator is the one used throughout this dissertation: even when referring to other sources I will convert their notation into my standardised form except in situations where the notational difference is significant.

Table 2.1 — Symbols Used for Logical Operators

Concept	Variety of Symbols Used
Negation	~ , - , ¬
Conjunction	& , . , ∧
Disjunction	v, +
Material implication	⇒ , → , ⊃
Bi-conditional	⇔ , ≡, ~
Existential quantifier	(∃x), (Ex)
Universal quantifier	(∀x) , (x)

Although many of these differences are purely cosmetic, seeing a variety of symbols in different textbooks is nonetheless confusing to students. Some symbols encourage an intuitive meaning (eg "&" is an obvious symbol for conjunction, and the quantifiers "∃" and "∀" connote "Exists" and "All") while others seem arbitrary and may even obscure the intended meaning.

The use of the horseshoe symbol "⊃" is particularly problematic. When one writes P⊃Q, it could easily be assumed by someone familiar with set notation that Q is claimed to be a subset of P. This is not only incorrect, but in fact the reverse of what is intended: P⊃Q means that whenever P is true, Q is also true, whereas the assumed set expression Q⊂P means that whenever Q is true, P is true. Historically, the horseshoe predated the set notation. In 1891 Peano used the notation "b C a" to represent "b is a *consequence* of a" and its inverted form "a Ɔ b" to indicate material implication[3]. This "Ɔ " was turned into "⊃" by later writers. Although it seems commonly

[2] [PRIO55] is the most recent book I have located which uses Lukasiewicz's notation. The main advantage of a prefix notation is that parentheses are not required. This syntactic simplification is what makes Lukasiewicz's notation appropriate for the games of "WFF 'N PROOF" (first invented by L.E.Allen in 1956 and still marketed by WFF 'N PROOF Publishers, 1490 South Boulevard, Ann Arbor, MI 48104-4699, USA). It has also been reported that blind students find prefix notation easier to work with (Ron Barnett <rbarnett@grits.valdosta.peachnet.edu> (pers. comm.)), presumably because an expression can be parsed without backtracking.

[3] Randall Dipert <dipert@cs.fredonia.edu> (author of articles on the history of logical notation in both the Encyclopedia Britannica and the Routledge Encyclopedia of Philosophy) and Martin Huehne <huehne@brouwer.informatik.uni-dortmund.de> pers. comm.

accepted that this was the origin of the horseshoe symbol, Quine claims that the horseshoe was used in 1816 by Gergonne [QUIN40], and this is supported by [KNEA62 p350].

Quine initially used the "$\supset$", but changed to "$\rightarrow$" in 1982 because it "is now widely used and is more suggestive" [QUIN50 p26]. The arrow (either "$\Rightarrow$" or "$\rightarrow$") appeared in [HILB28], though I have not been able to establish whether that was the original source.

Avoiding student confusion is more important in this dissertation than the claim of historical precedence, and since the method of teaching logic used here is founded on set theory, the "$\supset$" symbol will be dropped in favour of the arrow.

2.3 Methods for Proving or Deriving Sequents

Once logical statements have been written in some symbolic notation, a method is required whereby the relationships between logical expressions may be established. Most importantly, we wish to know how to establish whether one expression logically follows from another. The usual way to write this is as a sequent of the form $A_1,A_2,\ldots,A_n \vdash C$ or $A_1,A_2,\ldots,A_n \models C$, which are both claims that the assumptions $A_1,A_2,\ldots,A_n$ entail the conclusion C. The first expression, $A_1,A_2,\ldots,A_n \vdash C$, is a syntactic claim that in some appropriately defined formal system, the wff C can be derived from the hypotheses $A_1,A_2,\ldots,A_n$, whereas the second, $A_1,A_2,\ldots,A_n \models C$, is a semantic claim that it is valid to deduce the conclusion C from the assumptions $A_1,A_2,\ldots,A_n$. There is both an important distinction and an important connection between these which is described in the following section.

A sequent (in either the syntactic or semantic sense) is a claim of a certain relationship between the wffs on the left of the entailment sign and the single wff on the right. Different systems of logic propose different approaches to substantiating such a claim, and the purpose of this section is to describe and compare these varied approaches. We shall begin with a purely syntactic approach which lays a precise and formal foundation for the rest of the discussion, and then cover several semantic approaches which are much more suited to an introductory course in logic. In each case the approach is illustrated with an example: the sequent $P, \sim(P\&Q) \vdash \sim Q$ (or $P, \sim(P\&Q) \models \sim Q$, as appropriate). A full comparison would require many more examples, but the purpose here is to illustrate the methods rather than to criticise them and so a single example will suffice.

2.3.1 Axiomatic Derivations

In an axiomatic system, a set of symbols is defined, along with formation rules which define how those symbols may be combined (ie rules which generate well-formed formulae). Certain wffs are then defined to be axioms, and certain syntactic operations are defined as rules of inference. A theorem is any wff which is either an axiom or which can be derived from other theorems by means of the rules of inference. In axiomatic systems we are not directly concerned with either truth or validity, but rather derivability.

This approach to logic was pioneered by Frege, Russell, Whitehead, Hilbert and Bernays [KNEA62 pp524–538]. Significant work was also done by von Neumann and Church [CHUR44]. It is the most mathematically precise of all the methods discussed here, but also the most cumbersome. Derivations of theorems would be impossibly long if it were not for various meta-theorems which allow useful theorems to be proved without the need to construct the complete derivation. Axiomatic systems dictate mechanical manipulation of meaningless symbols, and require extremely abstract reasoning skills. For this reason, they are difficult to teach to students who are learning logic for the first time, and hence rarely used in introductory textbooks, though they do feature in higher level mathematical textbooks. Recent books which follow an axiomatic approach include [SHOE67, MASS70, HAMI78 and DOWS86].

In Massey's system [MASS70], wffs consist of sentential variables, parentheses and the operators ~ and ⇒. There are three Axioms —

Axiom 1: P⇒(Q⇒P)

Axiom 2: (P⇒(Q⇒R))⇒((P⇒Q)⇒(P⇒R))

Axiom 3: (~P⇒~Q)⇒(Q⇒P)

and two Rules of Inference —

Rule 1: From A and A⇒B one may infer B (*modus ponens*)

Rule 2: From A, one may infer the result of substituting a wff B for a sentential variable c throughout A (*substitution*)

This system is similar to that in [CHUR44] and [HAMI78]. The set of axioms and rules is kept as small as possible in order to simplify the meta-theoretical analysis, while still being powerful enough to generate all tautologies of propositional logic[4]. Although this system only allows the two operators ~ and ⇒, other operators may be introduced as abbreviations for wffs containing only these two operators. For instance, the wff A&B may be thought of as an abbreviation for ~(B⇒~(B⇒A)) — a truth table (see the following Section) will show that these two expressions are equivalent. Thus, the wff ~(P&Q) in the example sequent needs to be rendered as ~~(Q⇒~(Q⇒P)) in Massey's system.

In such a system, the concept of a sequent is secondary: the primary concept is the derivability of theorems. A sequent in the form $A_1,A_2,\ldots,A_n \vdash C$ claims that if the wffs $A_1,A_2,\ldots,A_n$ are hypothesised in addition to the Axioms, then the wff C is derivable. The Deduction Theorem is a meta-theorem which asserts that the claim $A_1,A_2,\ldots,A_n \vdash C$ is provable if and only if $A_1,A_2,\ldots,A_{n-1} \vdash A_n \Rightarrow C$ is provable [MASS70 p144, HAMI78 p32]. Repeated applications of this

[4] In 1917, Nicod showed that one axiom and one inference rule is sufficient [KNEA62 p526].

meta-theorem yield the sequent $\vdash A_1 \Rightarrow (A_2 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow C)\ldots))$. In other words, the sequent $A_1, A_2, \ldots, A_n \vdash C$ is provable exactly when the *corresponding conditional* $A_1 \Rightarrow (A_2 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow C)\ldots))$ can be derived from the Axioms and Inference Rules with *no* extra assumptions (ie when $A_1 \Rightarrow (A_2 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow C)\ldots))$ is a theorem).

The example sequent P, ~(P&Q) ⊢ ~Q can be represented in Massey's system as P, ~~(Q⇒~(Q⇒P)) ⊢ ~Q and this may be translated into the corresponding conditional P⇒(~~(Q⇒~(Q⇒P))⇒~Q). If we could show that the latter was a theorem of the axiomatic system, then we could say that the original sequent was provable. A derivation of this wff would be very long and difficult to construct, and nobody in their right mind would attempt it. Instead of taking the direct approach, it is better to take a look at the whole system and prove some further meta-theorems first. For instance, we could prove the principle of Double Negation: that any wff ~~P is a theorem if and only if P is also a theorem [MASS70 p148]. We could prove that it is possible to add an assumption to a sequent without changing its provability: that if $A_1, A_2, \ldots, A_n \vdash C$ is provable, then so is $A_1, A_2, \ldots, A_n, A_{n+1} \vdash C$. We could prove the principle of *reductio ad absurdum*: that if $A_1, A_2, \ldots, A_n \vdash C$ is provable and also $A_1, A_2, \ldots, A_n \vdash {\sim}C$ then so is $A_1, A_2, \ldots, A_{n-1} \vdash {\sim}A_n$ [MASS70 p 151]. Having proved such meta-theorems, we could proceed to prove the sequent P, ~~(Q⇒~(Q⇒P)) ⊢ ~Q as in Figure 2.1.

(1)	⊢ P⇒(Q⇒P)	Axiom 1
(2)	⊢ (P⇒(Q⇒P))⇒(Q⇒(P⇒(Q⇒P)))	Substitute P⇒(Q⇒P) for P in (1)
(3)	⊢ Q⇒(P⇒(Q⇒P))	Modus ponens from (1) and (2)
(4)	Q⊢ P⇒(Q⇒P)	Deduction Theorem from (3)
(5)	P,Q⊢ Q⇒P	Deduction Theorem from (4)
(6)	P,Q,~~(Q⇒~(Q⇒P)) ⊢ Q⇒P	Adding an assumption to (5)
(7)	~~(Q⇒~(Q⇒P)) ⊢ Q⇒~(Q⇒P)	Double Negation
(8)	P,~~(Q⇒~(Q⇒P)) ⊢ Q⇒~(Q⇒P)	Adding an assumption to (7)
(9)	P,Q,~~(Q⇒~(Q⇒P)) ⊢ Q⇒~(Q⇒P)	Adding an assumption to (8)
(10)	P,Q,~~(Q⇒~(Q⇒P)) ⊢ ~(Q⇒P)	Modus Ponens from (9)
(11)	P,~~(Q⇒~(Q⇒P)) ⊢ ~Q	Reductio ad absurdum from (6) and (10)

Figure 2.1 — Proof of P,~(P&Q) ⊢ ~Q in an Axiomatic System

Although this takes only eleven lines, it must be remembered that this is not a complete derivation in the sense defined in the first paragraph of this section. The various meta-theorems assure us that such a derivation is possible, but the complete derivation would be very much longer. For instance, a full justification of line (7) alone would require at least fourteen lines.

Although the discussion above focuses on an axiomatic system for propositional calculus, axioms can be added for quantification in such a way that the system can generate all logically valid wffs (in a sense which can be made precise) of FOPL [HAMI78 p71, MASS70 p414].

Everything so far has been defined syntactically, without regard to the meaning of the symbols. Nevertheless the symbols, formation rules, axioms and inference rules have all been carefully chosen so that the resulting theorems can be interpreted in a meaningful way. Prior to interpretation, an axiomatic system is no more than a syntactic game; after being appropriately interpreted the theorems are seen to be exactly the tautologies of propositional logic.

2.3.2 Truth Tables

Truth tables are the standard method of assigning truth values to propositional wffs, that is, of adding semantics to the syntactic system described above.

First, we view the symbols (such as P and Q) as propositional variables which represent certain propositions: for instance P may represent the claim that "All swans are white", Q may represent the claim that "The temperature is 15 degrees". Each such variable may be either true or false.

Second, we construct tables of truth values in which columns indicate the truth (T) or falsity (F) of one or more propositional wffs, and rows indicate each of the possible assignments of truth values to the propositional variables. Truth tables are used to define primitive logical operators and also to analyse compound wffs which are formed by combining a number of those primitive operators. Figure 2.2 shows the truth table which is the standard definition of conjunction.

P	Q	P & Q
T	T	T
T	F	F
F	T	F
F	F	F

Figure 2.2 — Truth Table Definition of Conjunction

In some situations, especially when constructing or analysing computer logic circuits, the symbols '1' and '0' are used instead of 'T' and 'F' respectively. In such cases, truth tables are sometimes drawn in a form corresponding to a Karnaugh Map (see Section 2.3.6.5) as shown in Figure 2.3.

&	0	1
0	0	0
1	0	1

Figure 2.3 — Alternate Truth Table Definition of Conjunction

Once the concept of truth has been introduced, we can designate each propositional wff as either <u>tautologous</u> (those which are always true), <u>contingent</u> (those which are sometimes true and sometimes false) or <u>inconsistent</u> (those which are always false). For instance, the wff $P \Rightarrow (\sim(P\&Q) \Rightarrow \sim Q)$ is shown to be a tautology by the truth table in Figure 2.4 since the values under the main operator (the left-most $\Rightarrow$) indicate the wff to be true in all four rows.

	P	Q	P	⇒	(~	(P&Q)	⇒	~Q)
1	T	T	T	**T**	F	TTT	T	FT
2	T	F	T	**T**	T	TFF	T	TF
3	F	T	F	**T**	T	FFT	F	FT
4	F	F	F	**T**	T	FFF	T	TF

Figure 2.4 — Truth Table for P⇒(~(P&Q) ⇒ ~Q)

We may also introduce the concepts of a sequent and sequent validity. The sequent $A_1, A_2, ..., A_n \models C$ claims that the assumptions (or premises) $A_1, A_2, ..., A_n$ *logically necessitate* the conclusion C. (Notice the use of the semantic entailment sign $\models$ rather than the syntactic sign $\vdash$ used previously.) Such a claim is said to be valid if it is impossible for all the assumptions to be true unless the conclusion is also true.

Truth tables may be used both to prove and to disprove the validity of propositional sequents, and this may be done in two ways. Given the general sequent $A_1, A_2, ..., A_n \models C$ we may construct a truth table with columns for each of $A_1, A_2, ..., A_n$ and C. The sequent is valid if and only if, for every assignment of truth values to its variables for which all of $A_1, A_2, ..., A_n$ take the value "true", C takes the value "true" also. [LEMM65 p75, HAMI78 p23, GALT90 p47, ALLE92 p38]

Thus, the sequent P, ~(P&Q) $\models$ ~Q may be analysed by truth table shown in Figure 2.5.

	P	Q	P	~	(P&Q)	~Q
1	T	T	T	F	T	F
2	T	F	T	T	F	T
3	F	T	F	T	F	F
4	F	F	F	T	F	T

Figure 2.5 — Validation of P, ~(P&Q) ∴ ~Q by Truth Table

The wff ~(P&Q) has two columns of truth values under it. This reflects a two stage evaluation: truth values for P&Q are calculated first; and then those values are negated to give the final truth values for ~(P&Q). Notice that row 2 is the only row in which both premises are true, and in that row the conclusion is also true. Hence the sequent is shown to be valid.

Alternatively, Lemmon proves a meta-theorem to the effect that $A_1, A_2, ..., A_n \models C$ is valid precisely when the *corresponding conditional* $A_1 \Rightarrow (A_2 \Rightarrow (... \Rightarrow (A_n \Rightarrow C)...))$ is a tautology, and this of course may be confirmed by checking the truth table of $A_1 \Rightarrow (A_2 \Rightarrow (... \Rightarrow (A_n \Rightarrow C)...))$ [LEMM65 p76]. (This meta-theorem relates closely to the Deduction Theorem mentioned earlier.) Using this meta-theorem, we can translate the sequent P, ~(P&Q) ∴ ~Q into the corresponding conditional P⇒(~(P&Q) ⇒ ~Q) and construct a truth table for that wff. Figure 2.4 shows this wff to be true in all four rows. Hence the wff is a tautology and by the meta-theorem above, the original sequent is valid.

The advantage of these methods is that they are mechanical and guaranteed to work. However, for sequents containing many variables they become cumbersome and they are not applicable to sequents containing quantifiers.

2.3.2.1 Relationship Between Syntactic Entailment and Semantic Entailment

In Section 2.3.1 *syntactic* entailment, $A_1,A_2,\ldots,A_n \vdash C$, was defined to mean that if the wffs $A_1,A_2,\ldots,A_n$ are hypothesised in addition to the Axioms, then the wff C is derivable. In Section 2.3.2 *semantic* entailment, $A_1,A_2,\ldots,A_n \models C$, was defined to mean that the assumptions $A_1,A_2,\ldots,A_n$ logically necessitate the conclusion C. In both cases we have seen that the sequent may be proved by reference to the corresponding conditional $A_1 \Rightarrow (A_2 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow C)\ldots))$. In the syntactic view, the important characteristic of the wff $A_1 \Rightarrow (A_2 \Rightarrow (\ldots \Rightarrow (A_n \Rightarrow C)\ldots))$ is whether or not it is a theorem, whereas in the semantic view, the important characteristic is whether or not it is a tautology.

Two important meta-theorems indicate the relationship between these concepts: the Validity Theorem asserts that every theorem is a tautology, and the Completeness Theorem asserts the converse, that every tautology is a theorem. Given an appropriately defined syntactic system, and an appropriate allocation of semantics, these meta-theorems amount to an equivalence between the two approaches. In view of this, and since the technical distinction between syntactic and semantic approaches is not an important one in an introductory logic course, I will avoid any future mention of the distinction in this dissertation. Instead of using either $\vdash$ or $\models$, I will use the more neutral symbol $\therefore$. In the context of an introductory logic course, this has the advantage of making use of a symbol with which the students are already familiar, and has the added advantage that it is much easier for my word processor!

2.3.3 Natural deduction

Whereas in the axiomatic methods valid formulae are derived from a sequence of axioms by means of a few forms of inference, in a natural deduction system assumptions are proposed from which a sequence of logical deductions are made. The aim of this is to produce "proofs" which more closely mirror the natural sequence of human deduction.

There are a number of natural deduction systems and these may be categorised into two groups which I shall exemplify using the system of Paulson [PAUL87] and Lemmon [LEMM65] respectively. The key distinction between these two categories is that whereas Lemmon's proofs are strictly linear, Paulson's are tree-structured.

Both of these types of system apply to the whole of FOPL. Every sequent which can be proved by these natural deduction systems is necessarily valid, and every valid sequent of FOPL can be proved within both systems. However, such natural deduction systems are unable, in general, to *disprove* an invalid sequent.

2.3.3.1 Lemmon

In Lemmon's system, each line of a proof must be justified by one of fourteen rules of derivation. In the proof of P, ~(P&Q) ∴ ~Q shown in Figure 2.6, the four columns list not only the wff being derived (third column), but also a line number (second column), the rule of derivation used to deduce this wff (fourth column), and a list of line numbers indicating the assumptions upon which the wff depends (first column).

1	(1)	P	A	(Assumption)
2	(2)	~(P&Q)	A	
3	(3)	Q	A	
1,3	(4)	P&Q	1,3 &I	(Conjunction-introduction)
1,2,3	(5)	(P&Q)&~(P&Q)	4,2 &I	
1,2	(6)	~Q	3,5 RAA	(Reductio ad absurdum)

Figure 2.6 — Validation of P,~(P&Q) ∴ ~Q in Lemmon's Natural Deduction System

Various books [eg ALLE92, POSP74] and computer programs (eg LemmonAid and Deriver Plus) have been written to use Lemmon's system, but it is just one system of many in this category. Other authors [eg COPI61, MATE65, GUTT71, MCCA81, GALT90] have adopted or invented systems which are similar in that they require a series of wffs in a proof to be listed one per line, along with a justification for each wff based on previous lines. Such systems vary in two main ways: the set of prescribed derivation rules; and the method of keeping track of assumptions.

One of the main benefits of Lemmon's system over other natural deduction systems is the explicit tracking of assumptions. Many other systems simply indicate when an assumption is discharged, but Lemmon (and [MATE65]) requires that a list of assumptions be maintained constantly. Because of this, at any point in the proof it is clear what has been proved so far. So in line 6 of Figure 2.6 it is clear that we have proved the sequent, since the wff ~Q has been derived based on the assumptions 1 (ie P) and 2 (ie ~(P&Q)).

There is a notable similarity between the structure of the Lemmon-style proof and the axiomatic proof in Section 2.5. However, in Lemmon, reductio ad absurdum is a primitive inference rule whose legitimacy is based on intuition. In contrast, reductio ad absurdum has no place inside Massey's axiomatic system, but is a meta-theorem. Any proof in Massey's system which uses reductio ad absurdum is actually an abbreviation for a longer (much longer!) derivation from the axioms.

2.3.3.2 Paulson

Paulson's system is based on the approach of Gentzen [GENT35] (see also [KNEA62 p538]). Paulson allows rules of assumption, contradiction, and both introduction and elimination rules for each of conjunction, disjunction, implication, negation, universal quantifier and existential quantifier. In this system, the sequent P, ~(P&Q) ∴ ~Q can be proved as shown in Figure 2.7. Note that apart from lines of assumption, each line is inferred from one or more previous lines. (The symbol $\wedge$ indicates a contradiction.)

$$\dfrac{\dfrac{\dfrac{\text{P}\therefore\text{P (Assumption)} \quad \text{Q}\therefore\text{Q (Assumption)}}{\text{P,Q}\therefore\text{P\&Q (\&-introduction)}} \quad \text{\textasciitilde(P\&Q)}\therefore\text{\textasciitilde(P\&Q) (Assumption)}}{\text{P,Q,\textasciitilde(P\&Q)}\therefore\wedge \text{ (\textasciitilde-elimination)}}}{\text{P,\textasciitilde(P\&Q)}\therefore\text{\textasciitilde Q (\textasciitilde-introduction)}}$$

Figure 2.7 — Validation of P,~(P&Q)∴~Q in a Gentzen-style Natural Deduction System

Gentzen proved that in such a system, every proof can be written in a normal form such that any derived formula appears only once. In a certain sense, these proofs are as direct as possible, with no sidetracks (ie inferences which do not lead towards the final conclusion). This may be seen as an advantage over a Lemmon-style system which does allow unnecessary sidetracks. However, Gentzen notes that natural human reasoning is necessarily linear [GENT35 p76] and so the linearity of Lemmon perhaps makes it more "natural".

2.3.4 Resolution

The method of resolution relies on three principles: first, that if the *negation* of a required conclusion is inconsistent with the premises, then it is legitimate to claim that the conclusion is entailed in the premises; secondly, that every wff may be expressed in conjunctive normal form (that is, a conjunction whose conjuncts are each elementary disjunctions); and thirdly, that if a term and its negation appear in separate clauses, they may be eliminated. (The sense of the terms "inconsistent", "conjunctive normal form" and "clause" are exemplified in Figure 2.8.)

This third principle is the Resolution Rule, which is the only rule of inference ever used in this system (excepting the various rules of equivalence which must be used to convert the original sequent into conjunctive normal form). The Resolution Rule allows the following inference —

$$\dfrac{\begin{array}{c} LvK_1vK_2v...vK_k \\ \sim LvM_1vM_2v...vM_m \end{array}}{\therefore\ K_1vK_2v...vK_kvM_1vM_2v...vM_m}$$

Applying these principles to the example P,~(P&Q)∴~Q we can construct the proof by resolution as shown in Figure 2.8.

Step 1: We wish to establish whether the two premises P and ~(P&Q) and the negation of the conclusion ~Q are mutually satisfiable. In other words, is P&~(P&Q)&Q satisfiable (in the sense exemplified below)?

Step 2: Convert this wff into conjunctive normal form, then list the conjuncts as a set of *clauses* . So P&~(P&Q)&Q becomes (P)&(~Pv~Q)&(Q), which may be written as —

C_1: P

C_2: ~Pv~Q

C_3: Q

Step 3: Progressive simplification is now performed using the Resolution Rule.

C_4: ~Q (From C_1 and C_2, since one contains the term P and the other contains the term ~P)

C_5: □ (From C_3 and C_4, since one contains the term Q and the other contains the term ~Q)

After two applications of the Resolution Rule, we are left with the empty clause (□), indicating that the original clauses (C_1,C_2,C_3) are unsatisfiable. Hence the wff P&~(P&Q)&Q is unsatisfiable and so the sequent P,~(P&Q)∴~Q must be valid.

Figure 2.8 — Validation of P,~(P&Q)∴~Q by Resolution

This method is far from straight-forward for humans, yet ideally suited to computers and hence Resolution forms the basis of much computational logic and automated theorem proving. Although the example above is very simple, the method can be applied to any sequent of FOPL. See [EISI93] for a detailed exposition. Various computing textbooks present this method, for instance [ROBI79, MANN85, DOWS86, GALT90, and AHO92]. Given its importance in automated theorem proving and artificial intelligence, it is appropriate to teach Resolution to Computer Science students, though because of its unnaturalness it is perhaps best to teach it in some course subsequent to a first introduction to logic.

2.3.5 Semantic tableaux

The early work on semantic tableaux was carried out by Beth [BETH59] and Smullyan [SMUL68] though they cite similar techniques by Hintikka (1955) and Schütte (1956). Some authors call this type of system "truth trees", eg [JEFF67].

A semantic tableau attempts to establish whether the premises of a sequent are consistent with the *negation* of the conclusion. If it is found to be so, then the original conclusion is not necessitated by the premises and so the sequent is invalid. Conversely, if the negation of the conclusion is found to be inconsistent with the premises then the sequent is valid. The method is a diagrammatic version of the Resolution strategy described above.

In the system described in [JEFF67], the premises are written one after the other, followed by the negation of the conclusion. Compound wffs are processed one at a time in a tree-like structure according to eight patterns (two for each connective &,v,⇒, and ⇔). Whenever a path through the tree contains some wff as well as the negation of that same wff, the path is closed (marked with an "x"). The sequent is valid if and only if all paths through the final tree are closed. (Any unclosed path in the final tree indicates a set of variable assignments which satisfy the initial set of wffs, which shows that the negation of the conclusion is not inconsistent with the premises.)

The sequent P,~(P&Q)∴~Q is shown to be valid by the semantic tableau in Figure 2.9. Lines 1 and 2 list the two premises and line 3 shows the *negation* of the conclusion[5]. We now look for any compound wff which has yet to be processed, and in this case we find only ~(P&Q). This pattern is processed by forming two new branches at the bottom of the tree and writing ~P after one branch and ~Q after the other (that is, line 4). Having done this, we place a tick at line 2 to indicate that the wff has been processed. Now we see that one path through the tree gives the sequence P, ~(P&Q), Q and ~P. Since this path contains a contradiction (both P and ~P) we close the path by placing an "x" at the bottom. Another path through the tree gives the sequence P, ~(P&Q), Q and ~Q. This also contains a contradiction (Q and ~Q) and is likewise closed. The process stops here because all paths through the tree have been closed (apart from the fact that there are no more compound wffs to be processed).The fact that there are no open paths through the completed tree indicates that the wffs at lines 1,2 and 3 are inconsistent, and hence that the sequent is valid.

(1)		P	
(2)	√	~(P&Q)	
(3)		Q	
(4)		~P	~Q
		x	x

Figure 2.9 — Validation of P,~(P&Q)∴~Q by Semantic Tableau

As Reeves and Clarke point out, the proofs which result from semantic tableaux are not a natural sequence of easily justifiable steps [REEV90 p89]. It is meaningless to pick some point in the middle of the proof and ask "what has been proved so far?" Even given a completed proof, it is impossible to translate the tableau into a natural English explanation. Nevertheless, the system is easy to learn (primarily because it is mechanical) and may be extended to encompass all FOPL.

5 In this system double negations are erased immediately.

Apart from sources mentioned above, other textbooks of logic which present semantic tableaux include [GUTT71, BELL77, and FISH88]. [ROBI79] presents an approach which combines Gentzen's style of natural deduction (see Section 2.3.3.2) with semantic tableaux.

2.3.6 Diagrammatic methods

Since this research presents a diagrammatic method for establishing the validity of propositional sequents, it is useful to describe other types of diagrammatic approaches to logic. We shall see, however, that none of these other diagrammatic tools are directly applicable to the analysis of propositional sequents.

2.3.6.1 Euler's Circles

When dealing with categorical statements, it is useful to illustrate relationships between categories with some diagram. This idea may have been initiated by Leibniz, but was popularised by Euler in 1761 [KNEA62 p349]. Euler's Circles were an early method of representing categorical relationships diagrammatically. The four types of categorical statement are represented by the four separate diagrams shown in Figure 2.10.

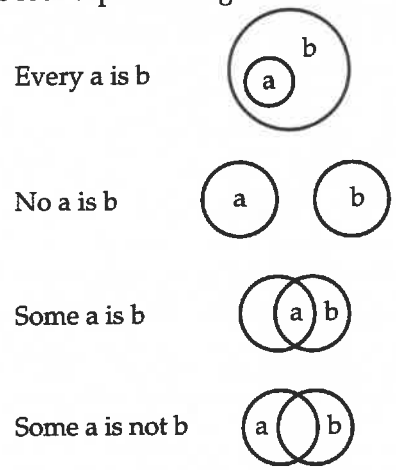


Figure 2.10 — Euler Circles for Categorical Statements

Underlying these Circles are the traditional assumptions that no category is empty and that "some" means "at least one, but not all". Modern logic tends not to make these assumptions.

Since Euler's Circles are intended for the analysis of categorical statements, it is difficult to use them to represent the example sequent P,~(P&Q)∴~Q. Rather than force an unnatural correspondence between our modern notation and the structure of Euler's Circles, we shall omit the example.

Textbooks of logic which use Euler Circles include [WELT11] and [JOYC56].

2.3.6.2 Venn Diagrams

Rather than use separate diagrams for each type of categorical statement, Venn combined all cases into one diagram [VENN1881]. Thus, two interlocking circles can represent any categorical relationship between two categories. The system can even be extended to relationships between more than two categories: three overlapping circles may easily be drawn; four with some difficulty; larger numbers are impractical. By marking different areas of the diagram (perhaps by shading), Venn's approach is flexible enough to diagram not only categorical statements, but other logical expressions as well. Venn Diagrams are frequently used to teach set concepts, and set terminology is often used to explain how logical expressions are represented on a Venn Diagram.

Propositional expressions are not immediately amenable to representation on a Venn Diagram since propositions are not the same as categories. Nevertheless, with a slight shift in interpretation we can proceed as follows. Instead of thinking of a proposition P as meaning "P is true", think of it as meaning "the collection of P-things" or "those things which satisfy P". Then the propositional expression ~(P&Q) can be thought of as meaning "it is not the case that there are things satisfying both P and Q". Alternately, one could interpret P as the set of all possible worlds in which P is true.

An analysis of the sequent P,~(P&Q)∴~Q is shown in Figure 2.11. In this Figure, U represents the universe of discourse, shading is used to show the regions which represent the two premises P and ~(P&Q), and the intersection of those two regions is doubly shaded. Thus the combined premises are represented by the left lobe of the circle labeled "P". After drawing such a diagram to represent the premises of a sequent, one can then consider the conclusion. In this case, we see that the left lobe of the circle P is completely outside the circle Q. Hence any situation which satisfies the premises must also satisfy ~Q. Another way of expressing this is to say that the region representing the premises is a subset of the region representing the conclusion. Thus we find that P,~(P&Q)∴~Q is valid.

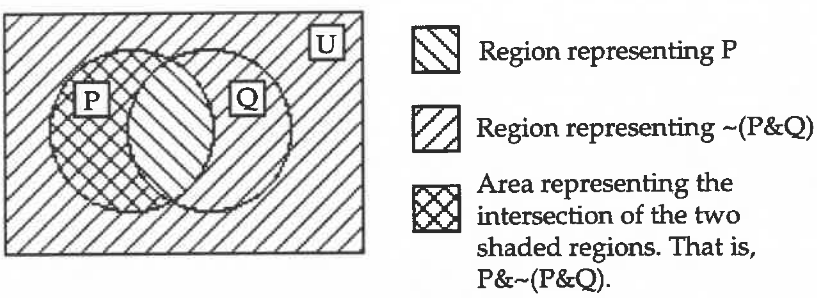


Figure 2.11 — Validation of P,~(P&Q)∴~Q by Venn Diagram

It has been noted that Euler's Circles assume that no category is empty and that "some" means "at least one, but not all". Venn Diagrams *may* be viewed with these assumptions, but may also be interpreted in the modern way, since any region of a Venn Diagram may be empty.

Venn Diagrams are much more common than Euler Circles and have featured in many textbooks over the last 110 years, including [KEYN1884, CARR1896, MELL04, QUIN50, PRIO55, KNEE63, KORF74, POSP76, HOCU79, GEAC76, THEW83, DAVI86, and AHO92]. Venn Diagrams are convenient for simple examples and provide a very clear visual tool for teaching basic logical principles. They can be used to assist in the analysis of propositional expressions (provided there are fewer than five variables) and those expressions of predicate logic which permit a categorical interpretation. However, they are not useful for more complex expressions of FOPL.

2.3.6.3 Carroll Diagrams

Lewis Carroll also devised a method of diagramming logical statements [CARR1896]. This method is similar to both Venn Diagrams and to Karnaugh Maps (see Section 2.3.6.5). Carroll himself provides a comparison of various methods of solving categorical syllogisms in [CARR1896 pp246–249].

A Carroll Diagram for two variables is simply a square split into four quarters as shown in Figure 2.12a. The North Half (to use Carroll's terminology) is assigned to the first variable x; the South Half to not-x; the West Half to y; and the East Half to not-y. These four *Cells* play the same role as the four areas in a Venn Diagram (Figure 2.11). When any Cell in a Carroll Diagram is known to be empty, it is marked with a 'O'; when it is known that a particular Cell contains at least one Thing, it is marked with an 'I'. The Carroll Diagram for three variables is shown in Figure 2.12b, and Carroll describes an ad-hoc scheme for constructing diagrams with up to ten variables in [CARR1896 pp244–246].

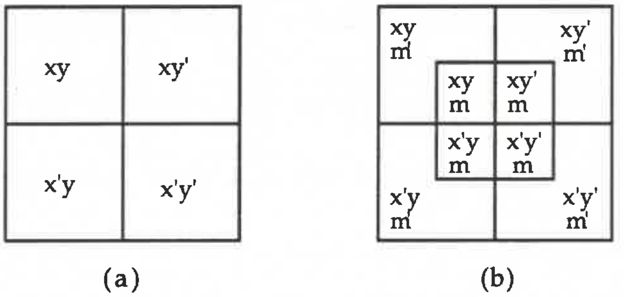


Figure 2.12 — Carroll Diagrams for Two and Three Variables

The procedure for analysing arguments using Carroll Diagrams is roughly the same as for Venn Diagrams. When applied to non-categorical expressions, the method is as uncomfortable as it is for Venn Diagrams and we shall not attempt it here.

Carroll Diagrams are rarely seen any more, though they are used in [GEAC76].

2.3.6.4 Lambert Lines

Inspired by J.H.Lambert (1728–1777), Binkley[6] has developed a method of diagramming syllogistic logic which is similar to Venn Diagrams. The key novelty of Binkley's method is that the diagrams can easily be typed using the standard 7-bit ASCII character set and thus can be readily stored and transmitted via computer networks.

Binkley uses horizontal lines to indicate that there is at least one thing which meets a certain condition[7]; a row of full-stops to indicate that they may or may not be anything meeting a certain condition; and an empty space to indicate that it is known that there are no things which met a certain condition. Using this notation, the four types of categorical proposition may be represented in Lambert Lines as shown in Figure 2.13.

```
All a are b              a: ======......
                         b: ============
No a are b               a: ======
                         b:       ======
Some a are b             a:    ===......
                         b: ...===
Some a are not b         a: ......===
                         b: ======
```

Figure 2.13 — Lambert Lines for Categorical Statements

Once again we note that the sequent being used as an example is not expressed in the form of a categorical syllogism and so does not fall within the intention of Lambert Lines. The example *could* be forced into the required mould, but this is somewhat artificial.

In summary, the diagrams of Euler, Venn, Carroll and Binkley are very useful for representing and analysing categorical statements (which was the purpose of their design) but not for representing more complex logical expressions.

2.3.6.5 Karnaugh Maps

Karnaugh maps are commonly used as an aid in the design of digital circuits. A brief introduction to boolean algebra is required before the role of Karnaugh maps can be appreciated —

A boolean expression is an algebraic expression in which the variables may take only the values 0 or 1, and in which the algebraic operations are negation (indicated by a prime), disjunction (indicated by addition) and conjunction (indicated by multiplication). For instance, the

---

[6] Robert Binkley <rbinkley@julian.uwo.ca> (pers. comm.), who lectures at the University of Western Ontario, Canada.

[7] In fact there are two different types of horizontal line: one for the traditional interpretation that categories are never empty, and one for the modern interpretation in which such an assumption is unnecessary.

propositional wff (P&Q)v(P&~Q) may be written as the boolean expression P.Q+P.Q′. Boolean expressions may be manipulated algebraically: eg P.Q+P.Q′=P(Q+Q′)=P(1)=P. This example shows four different boolean expressions which are equivalent, although the last (P) is clearly the simplest.

An important task in digital circuit design is to *minimise* a boolean expression: that is, to find the simplest equivalent boolean expression. Thus P.Q+P.Q′ may be minimised to P. A Karnaugh map is a rectangular grid which is a useful tool for this task of minimalisation. Up to four variables may be represented in just one rectangular grid, but more than four variables require multiple grids and are not easy to visualise. Once a Karnaugh map has been constructed for a boolean expression, it is easy to rewrite the expression as a sum-of-products (Disjunctive Normal Form) and to construct the corresponding digital circuit.

A Karnaugh map shares some similarities with Carroll diagrams; in fact the two are identical for expressions with only two distinct variables. While such diagrams are very useful for the task of minimalising a single boolean expression, they offer little help for the task of analysing propositional sequents. A propositional sequent (eg P, ~(P&Q) ∴ ~Q) would have to be converted into its corresponding conditional (eg P⇒(~(P&Q)⇒~Q)), then expressed using only negation, disjunction and conjunction (eg ~Pv((P&Q)v~Q))), and then translated into algebraic notation (eg P′+P.Q+Q′) before it could be displayed on a Karnaugh map. The Karnaugh map would show whether or not there was a simpler expression equivalent to the original sequent. In the case of a valid sequent, the Karnaugh map would show that the simplest equivalent expression was 1.

Many textbooks for computer science describe Karnaugh maps [eg KORF74, BART60, THEW83, DOWS86 and AHO92], but always in the context of minimalisation of boolean expressions rather than the context of propositional calculus.

2.3.6.6 Possible Worlds Diagrams

Bradley and Swartz introduced what they call "worlds-diagrams" in [BRAD79]. The status of a propositional expression in all possible worlds can be represented by a series of rectangular diagrams. An expression containing just one variable requires three diagrams; an expression with two distinct variables requires fifteen diagrams; an expression with five distinct variables requires 4,294,967,295 diagrams! Worlds-diagrams help students to understand the complete range of possible worlds to which a wff may applied, and illustrate the basic logical operations (such as conjunction, disjunction, negation, implication, quantification, necessity, and possibility) within those worlds. They have some advantages when working with modal operators, but are more cumbersome than truth tables for truth-functional logic, and are not practically useful when analysing compound expressions of FOPL.

Chapter 3: Typical Student Difficulties with Formal Logic

3.1 Introduction

Logic of any sort is not always natural to the uninitiated, but formal symbolic logic is the least intuitive and hence the most difficult to learn. Various forms of induction and informal logic flow naturally from people's experiences and teaching such forms of logic is largely a process of refining the students' pre-existing reasoning skills. However, when a rigid symbolism is introduced, coupled with equally rigid rules for manipulating that symbolism, the student is faced with a form of reasoning which can seem totally unlike their natural forms of inference and even contrary to that reasoning. (See [GILH88] for a well-balanced review of research into the extent to which people use logic in their normal thought patterns, especially Chapter 5.)

The very possibility of teaching reasoning skills has been doubted by many. See [NISB87] for a synopsis of the views of Thorndike (that problem solving is domain-specific and that logical principles taught in one domain are not transferable) and Piaget (that the learning of inferential rules depends on spontaneous cognitive development resulting from active self-discovery rather than on explicit instruction). Under these views, formal training in logic can never be effective and it would be difficult to make any claims about which methods of teaching logic are better than others.

Fortunately there *is* evidence to the contrary. Nisbett et. al. argue that rules of logical inference *can* be taught to the extent that they correspond with pre-existing pragmatic reasoning schemas [NISB87]. These conclusions are based on results reported in [CHEN85] and [CHEN86] and some other unpublished data.

In this dissertation I accept as a basic assumption that logic can be taught. Such an assumption is well justified in [NICK85]. However, it is clear that logic is not *easily* taught and therefore the search for effective methods of teaching logic is meaningful. This chapter examines some of the difficulties which students encounter when trying to learn a system of formal logic, and suggests ways in which these difficulties may be either avoided or overcome.

3.2 Difficulties with Rigid, Structured, Abstract Thinking

Mathematics, logic and computer science all require the ability to reason with formal rules. The concepts and procedures required in mathematical proofs, numerical calculations, manipulation of symbolic logical expressions, and computer programming are typically rigid, highly structured and abstract. Students must develop the ability to set aside intuitions and force themselves to adhere to the rules. Perhaps this is most clear in the case of debugging a computer program.

When a program fails to run as intended by the programmer, the programmer attempts to step through the program to discover the point at which the computer acted differently from the programmer's intention. This process of locating a bug requires that programmers put themselves in the place of the computer and follow exactly the rules by which the computer operates.

This ability to "think like a computer" is difficult for first year computer science students, perhaps because many may not have passed from Piaget's Concrete Operations stage to the Formal Operations stage (see [NICK85 p32] and further discussion in Section 6.5.6). Thus, they may be able to understand and use concrete procedures, but not abstract ones. This difficulty affects not only students' ability to learn computer programming but also their ability to learn formal logic. When teaching the course on logic described in Chapter 6 and the Appendix, I have consistently found that the more formal sections (especially the introduction of axiomatic systems) always cause students the most trauma.

Students also find it difficult to take a logical problem described in English, extract the significant logical features and translate them into symbolic form. The relationship between expressions and operations in an informally described logical problem and equivalent expressions and operations in a formal system is not at all trivial. Notwithstanding the comments in the previous paragraph, the actual formal manipulation of logical expressions is often much easier for the students than the initial task of formulating the problem appropriately.

One should not imagine that a student who has mastered certain formal techniques for the manipulation of logical expressions will necessarily be able to apply those techniques to a real-world problem. Even when a student has learnt to think in a rigid, structured and abstract way, the connection between the formal and the practical is not automatically apparent. Section 6.5 describes a number of educational principles which seek to address these difficulties.

3.3 Difficulties with the Truth-Functional Definition of Material Implication

Material implication is an attempt to capture the essence of conditional statements: that is (in English at least), statements of the form "if ... then ...". However, statements of this form have a variety of intentions, not all of which are truth-functional. Thus, no truth-functional definition of implication will be able to capture fully the diversity of meanings in conditional statements. Nevertheless, the truth table in Figure 3.1 represents the commonly accepted truth-function for material implication.

P	Q	P⇒Q
T	T	T
T	F	F
F	T	T
F	F	T

Figure 3.1 — Truth Table for Material Implication

Students' inability to apply this definition springs from two sources: first, people tend to reason incorrectly about conditional statements; and second, the above truth table violates their intuition about conditional statements with false antecedents. The following sections expand on these problems and examine a variety of proposed solutions.

3.3.1 O'Brien and Shapiro

In a series of studies, O'Brien found that less than 10% of the subjects reasoned correctly about conditional statements [OBRI72, OBRI73, SHAP73]. Subjects for these studies included late high-school students, college students and medical students.

O'Brien identified four forms of reasoning relative to conditional statements, the first two valid and the latter two invalid:

1. Modus ponens: P⇒Q, P ∴ Q
2. Contrapositive (modus tollendo tollens in this thesis): P⇒Q, ~Q ∴ ~P
3. Inverse (denying the antecedent): P⇒Q, ~P ∴ ~Q
4. Converse (affirming the consequent): P⇒Q, Q ∴ P

These four forms are listed in order of increasing difficulty. That is, when subjects were presented with English examples of each of the four forms, they showed significantly higher error rates for the second form when compared with the first form, for the third form when compared with the second form, and for the fourth form when compared with the third form.

From each subject's responses, O'Brien and Shapiro attempted to infer the subject's mental model of a conditional statement. The majority of subjects seemed to equate P⇒Q with (P&Q)v(~P&~Q). A large group seemed to think that P⇒Q, P ∴ Q and P⇒Q, Q ∴ P were valid forms and that all other forms were undecideable. Both of these interpretations were more common that the correct interpretation.

3.3.2 Wason's Four-Card Selection Problem

Since 1966 many variations of Wason's Four-Card Selection Problem have been used to substantiate the claim that the majority or people reason incorrectly about conditional statements [GILH88 pp113–123, EVAN82 Chapter 9]. In Wason's original experiment, the subjects were presented with four cards lying on a table, showing values similar to Figure 3.2.

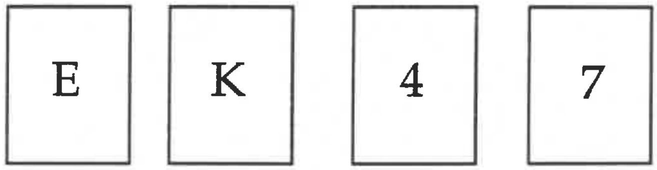


Figure 3.2 — The Original Four-Card Selection Problem

The task required of the subjects was to indicate which cards they must turn over in order to test the claim that "If a card has a vowel on one side then it has an even number on the other side". The correct answer is to choose those cards which have the potential of falsifying the claim, that is the card showing 'E' and the card showing '7'. This is clear if one views the claim in the form P⇒Q (where P stands for "this card has a vowel on one side" and Q for "this card has an even number on one side"), and considers the combinations of truth values for P and Q for which P⇒Q is false. However, only between 4% and 10% of subjects choose this combination. By far the majority of subjects chose the 'E' and the '4'.

The experiment has been repeated by many researchers, with a wide variety of subjects (including members of MENSA [WASO83 pp54–56]), and innumerable variations in experimental design, but the results remain consistent. The task seems simple enough, but the vast majority of people do not solve it correctly.

It is not required here to discuss the various attempts to explain these results, but simply to indicate that conditional statements are difficult for most people to master. It has been claimed that incorrect answers to this task arise from faulty internal truth tables, and that if subjects were first made aware of the correct truth table for material implication then the mistakes would be avoided. However, Wason denies this, claiming that the common mistakes in subjects' mental truth tables do not explain why they would choose '4' rather than '7', and adds that —

> Our experience suggests that a lecture on truth tables, an awareness of the possibilities which could occur on the other side of the cards, or exposure to conditional sentences in a more familiar form, would not be helpful. [WASO83 p48][1]

3.3.3 Teaching Experience

My own experience with teaching symbolic logic is that when students encounter the truth-functional definition of material implication (Figure 3.1), they accept the first and second lines

---

1 This comment highlights the role of transfer: even when subjects have learnt propositional calculus, they fail to transfer that knowledge to this selection task. See further comments on the importance of transfer in Section 6.5.3.

of the truth table, but react strongly against the third and fourth. Not only do they seem confused, but they actively argue against them. This experience is common to all teachers of logic with whom I have communicated.

To the student, it seems ludicrous to suggest that when a statement of the form "if ...then ..." has a false antecedent, then the overall statement should be considered to be true. The truth table for material implication is not at all self-evident and must be accompanied by some explanation.

3.3.4 Approaches to Explaining Material Implication

Given the difficulties students encounter with conditional statements, and with understanding the definition of material implication as a truth function, it is interesting to compare the methods used to justify this definition. There are many approaches to this: some indicate a primary concern with technical precision while others show varying degrees of concern for avoiding or alleviating student distress.

3.3.4.1 Definition by Truth Table

Most commonly, material implication is defined by truth table (Figure 3.1) or some verbal equivalent such as "X→Y is always true if X is false and also if Y is true" [HILB28 p4] or "A conditional sentence is false if the antecedent is true and the consequent is false; otherwise it is true" [SUPP57 p6].

This definition is justified by the authors in various ways, though frequently no justification is given at all [HILB28, COOL42, WERK48, BELL77, ROBI79, MANN85[2], PAUL87, REEV90]. Quine claims that the given truth table "constitutes the nearest truth-functional approximation to the conditional of ordinary discourse" [QUIN40 p15] and adds that this definition dates back to Philo of Megara[3]. [AHO92], [DOWS86] and [BASS53] take the same approach, admitting that this truth function does not always match English usage. Suppes takes a bold approach and simply states that in maths and logic this is the way it is done! [SUPP57]

Shoenfield defines material implication as a function rather than as a truth table, though the effect is the same. He claims that this definition follows the "mathematical meaning of if ... then" [SHOE67 p11].

---

[2] Manna and Waldinger have the added novelty of defining a truth table for "if ... then ... else"! [MANN85, p13]

[3] Peirce provides a summary of the debate between Philo and Diodorus on whether hypothetical propositions (ie conditionals) are at all different from categorical propositions. Philo claims (and Peirce agrees) that the forms "If P then Q" and "Every P is Q" are identical, but Diodorus (supported by Peirce's contemporary Shröder) claims they have different meanings. [HART74, paragraph 3.439ff, written in 1896]

A variant on the truth table approach, shown in Figure 3.3, starts by showing that sixteen distinct truth tables may be constructed for two variables. After columns 1 (tautology), 16 (inconsistency), 2 (disjunction), 8 (conjunction) are discussed and named, the author draws the reader's attention to column 5 and says in effect "this is an interesting and useful column so let's give it a name as well". This approach is taken by [JEFF67 p49] and [KORF74 p254]. Jeffrey also comments "Except in odd cases the truth conditions for the indicative English conditional are accurately given by the usual truth table [ie Column 5 in Figure 3.3]" [JEFF67 pviii].

P Q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
T T	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
T F	T	T	T	T	F	F	F	F	T	T	T	T	F	F	F	F
F T	T	T	F	F	T	T	F	F	T	T	F	F	T	T	F	F
F F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F

Figure 3.3 — Sixteen Possible Truth Tables

3.3.4.2 Definition in Terms of Other Operators

Other authors define material implication as an abbreviation of some other boolean expression. The normal form of this definition is (P⇒Q) $=_{def}$ (~PvQ) [WHIT10, STEB30, STEB43, EATO31] while others use (P⇒Q) $=_{def}$ ~(P&~Q) [QUIN41, COPI61, MITC62, CARN64, KIRW78, HOCU79]. Virtually all books show these equivalences at some point. Several texts explicitly note that the two definitions are interchangeable, and Ambrose and Lazerowitz make a major point of showing that not only can material implication be defined in terms of negation and conjunction, but one could equally well define material implication as the primitive operation and then define disjunction and conjunction in terms of negation and material implication[4] [AMBR48]. They prefer the definition (P⇒Q) $=_{def}$ ~(P&~Q) over (P⇒Q) $=_{def}$ (~PvQ), saying that while the first makes good English sense, the second looks problematic, even though the two are provably equivalent [AMBR48 p75].

When Davis introduces the horseshoe operator he defines it to be equivalent to both ~(P&~Q) and (~PvQ). Long before this, however, he discusses non-truth-functional conditional statements (using the symbols ⇒ and →) and presents the idea of material implication in syllogistic arguments. [DAVI86]

Like the explicit definition by truth table, this form of definition is sometimes not accompanied by any justification [WHIT10, STEB30]. In her later work, Stebbing notes that what ~PvQ defines is *material implication*, which is not necessarily the same as the English "if ... then ..." structure [STEB43 p139].

---

[4] There is at least one book which actually takes this approach: [BELL77].

3.3.4.3 Definitions Relying on Examples

Several authors use selected examples (either in mathematics, science or conversational English) to justify their definition of material implication. Rosser uses some verbal trickery to show that English phrases of the form "if A then B" are the same as "we cannot have both 'A' true and 'B' false" [ROSS53, p15]. He also cites a number of mathematical examples.

Massey uses the example that "If my memory is correct, then I owe you a dollar" means the same as "Either it is false that my memory is correct or else I owe you a dollar", and using this example shows that P⇒Q is simply an abbreviation for ~PvQ [MASS70 p52].

[GUTT71], [KELL90], [SAIN91] and [PRIO55] all take similar approaches.

Hermes claims that the structure ~(P&Q) occurs frequently in mathematics and that mathematicians express this as "if P then Q". With this as justification, he defines the two to be equivalent [HERM73]. Hamilton, writing specifically for mathematicians, justifies his truth table definition with the example "if n>2 then n^2>4" which, he says, is still a true statement even when *n* happens to be less than two [HAMI78 p5].

Quine draws on an English example to convince the reader of his claim that "if P then Q" is equivalent to ~(P&~Q), though he does this without explicitly constructing truth tables [QUIN41 p20]. In a later work, Quine writes —

> An affirmation of the form "if P then Q" is commonly felt less as an affirmation of a conditional than as a conditional affirmation of the consequent. If, after we have made such an affirmation, the antecedent turns out true, then we consider ourselves committed to the consequent, and are ready to acknowledge error if it proves false. If on the other hand the antecedent turns out to have been false, our conditional affirmation is as if it had never been made. [QUIN50 p21]

Consequently, he claims that the choice of declaring a conditional to be true whenever the antecedent is false is arbitrary. Kneebone, who relies on truth tables to define material implication rather than examples, makes a similar point —

> The truth-values that are to be ascribed to $\phi \rightarrow \varphi$ in cases in which ϕ is false are unimportant, since we do not draw conclusions from premises unless these are known to be true, or at least assumed to be true for the sake of the argument; but it greatly simplifies the formal logic of propositions if we define the truth-value of $\phi \rightarrow \varphi$ in all cases, taking it as T whenever ϕ has the truth-value F (compare with such conventional definitions in mathematics as a^0=1 and 0!=1). [KNEE63 p31, emphasis mine]

Mendelson also claims that the truth-functional definition is simply a convention, though it is also justified by the desire that (P&Q)⇒P should be a tautology [MEND64 p13].

Korfhage uses a pedagogically fascinating, though technically misguided analogy with computer programming to show why a conditional should be treated as true if the antecedent is false. He writes that given P⇒Q, where P is known to be false, it is true that we can't deduce anything about Q, however, we don't want the argument to stop there. Compare this with a FORTRAN program containing the statement "IF ALPHA .GT. X+7 GOTO 13". Even if "ALPHA .GT. X+7" is false, the overall statement is a good piece of FORTRAN and we want the program to continue running. Hence, so as not to disrupt an argument, we assign P⇒Q the value true whenever P is false. [KORF74]

This approach may appeal to students who already understand the FORTRAN "if" statement. However, there is some evidence that children who have been taught the "if...then...else" programming construct misconstrue conditional statements as bi-conditionals [SEID89]. That is, after being exposed to the program language interpretation of an "if" statement, they infer an incorrect truth table for material implication. Korfhage's approach is technically misguided since it confuses a form of conditional in which the antecedent and consequent are causally connected ("if the condition ALPHA .GT. X+7 is true then the next thing to do is execute the instruction at line 13") with the truth functional form which requires no causal connectivity.

The method of choosing or contriving an example which fits the author's intention is rather artificial. A more sophisticated approach is to note that the English "if ... then ..." structure is used in a variety of senses and that since we need to use symbolic operators unambiguously, we must choose just one of those senses. According to Church, "we select the one use of the words 'if ... then' ... in which they may be construed as denoting a relation between truth-values" [CHUR44 p38].

Reichenbach, who uses a truth table to define the horseshoe operator, makes a useful distinction between *adjunctive* implication and *connective* implication.

> It recently happened in Los Angeles that, while the screen of a movie theatre was showing a blasting of lumber jammed in a river, an earthquake shook the theatre. The implication "the blasting of lumber on the screen implied the shaking of the theatre" was then true in the adjunctive sense whereas it was false in the connective interpretation. ... We realise that the word "implies" here has not the same meaning as in conversational language; the implication in this case simply *adjoins* one statement to the other without *connecting* the statements. Adjunctive implication has a wider meaning than connective implication; if a connective

> implication holds, there also exists an adjunctive implication, but not vice versa. [REIC47 pp29,30]

Copi's "Introduction to Logic" has the most extensive variant of this approach. After explaining how we choose one of the two senses of the English "or" (the inclusive rather than exclusive sense), he lists four senses of implication and then chooses the one which can be written symbolically as ~(P&~Q). Choosing this interpretation over the other three is not arbitrary. Rather, he shows that this definition specifies the *common ground* between the four senses. [COPI61 pp245–252]

Galton argues that the truth table definition is the minimal truth-functional definition which will apply to all conditional statements.

> Even though → may not capture everything that is implied by "if ... then ...", at least we can say that a statement of the form A→B will be true whenever (if not more often than) "if A then B" is true, so that the English inference with statements of this form amongst its premises will be valid so long as the propositional calculus translation is. [GALT90 p57]

Georgacarakos and Smith devote many pages to this same point [GEOR79 p53ff].

3.3.4.4 Avoiding any Truth-Functional Definition

Another way to tackle the problematic definition of material implication is to avoid explicit definition altogether [FITC52, LEMM65, POSP74]. Fitch uses a natural deduction system in which conditional expressions may be manipulated by Modus Ponens and Distribution[5]. No mention is made of truth tables and (~PvQ)⇒(P⇒Q) is left as an exercise for the reader [FITC52]. Lemmon follows a similar path using the symbol ⇒ in proofs long before defining it as a truth function. He uses a natural deduction system with ten Rules of Derivation to prove that P⇒Q, ~PvQ and ~(P&~Q) are all interderivable. When he eventually gets around to discussing truth tables, it is then clear that P⇒Q should be defined to have the same truth table as both ~PvQ and ~(P&~Q). Even so, he admits that the truth table definition of material implication "seems rather arbitrary" [LEMM65 pp67–68].

3.3.4.5 Using Peirce's Notation[6]

Another approach relates material implication to the mathematical concept of less-than-or-equal-to. This is inspired by Peirce, who used a modified "≤" sign to stand for material implication in 1885 [HART74 paragraph 3.373]. The implication P⇒Q can be explained by showing that the truth of Q is at least as certain as the truth of P, because Q must be true

---

[5] The rule of Distribution may be symbolised as P⇒(Q⇒R) ∴ (P⇒Q)⇒(P⇒R).
[6] Suggested by John Sowa <sowa@turing.pacss.binghamton.edu> pers.comm.

whenever P is true, but Q can also be true on other grounds independent of P. If the value "true" is interpreted to be greater than the value "false", then the truth table for material implication will be identical to the truth table for less-than-or-equal-to.

This explanation may be especially useful for computer science students since they will already have a mental correspondence between "true" and "false" and the binary values 0 and 1.

3.3.4.6 Definition Based on the Idea of Contracts

In [NISB87], Nisbett et.al. claim that an effective way to teach conditional logic is to draw on pre-existing concepts rather than to define an entirely new concept. The pre-existing concepts they suggest are those of permission and obligation, both of which are forms of contract.

The statement "In order to do some action P you must have permission Q" follows precisely the same truth-function as $P \Rightarrow Q$. A contract of permission is violated only when the action P is performed without the required permission Q. If P is performed *with* permission Q, or if P is not performed at all, then the permission contract stands unviolated.

The statement "If you perform some action P then you are obligated to do Q" follows the same pattern. Such an obligation is violated only when P occurs but not Q, and hence an obligation schema behaves the same as material implication.

Note that permission and obligation are not presented simply as examples as described in Section 3.3.4.3. Rather, the aim is to proffer these to students as inference schema with which they are already well acquainted, and to indicate that processing conditional statements should be undertaken using those same schema.

3.3.4.7 Definition Based on Set Concepts

As can be seen in Chapter 5 and more expansively in the Appendix (pp A27ff), my own approach is quite different than any of those described above. If one can assume that students have an understanding of the basic concepts of set theory, then those concepts can be matched with parallel concepts in boolean logic. Negation can be explained as the logical counterpart to set complementation; disjunction as the counterpart to union; conjunction as the counterpart to intersection; material implication as the counterpart to the subset relation; and the bi-conditional as the counterpart to set equivalence.

The soundness of this as a teaching approach is indicated by these inter-related features:

- Since students already have a grounding in set theory, the approach defines new concepts in terms of familiar concepts;
- The visual model of set structures given by Venn Diagrams can immediately be transferred as a tool to aid understanding of logical expressions;

- By grounding the definitions of logical operations in set theory rather than by attempting to symbolise English statements, ambiguity is avoided[7]. For instance, the confusion about whether "or" is inclusive or exclusive disappears since a Venn Diagram makes it clear that union (and hence disjunction) is inclusive; and
- The definition of material implication is no longer "arbitrary", as some of the authors quoted above apologetically assert. The subset relationship $P \subseteq Q$ unambiguously disallows the situation where membership of P is true but membership of Q is false, and allows the other three possible situations. And so the definition of the logical counterpart $P \Rightarrow Q$ follows inexorably.

3.4 Difficulties with the Truth-Functional Definition of Disjunction

Errors in conditional reasoning have elicited much research and speculation, but it is not the only aspect of natural human reasoning which fails to fit nicely into a truth-functional model. Disjunctive reasoning has also been studied in this regard, and it is clear that the truth-functional definition given in Figure 3.4 does not always match people's intuitions.

P	Q	P v Q
T	T	T
T	F	T
F	T	T
F	F	F

Figure 3.4 — Truth Table for Disjunction

See [NEWS83 and EVAN82] for a summary of research into linguistic and conceptual factors related to difficulties with disjunctive reasoning. These references also describe Wason's THOG task, which is in some ways the disjunctive equivalent to the Four Card Selection Problem described above in Section 3.4.2.

The key difficulty with a truth-functional approach to disjunction is that in English the word "or" should sometimes be taken as inclusive but at other times as exclusive. For example, if entrance to a movie is restricted to people who are "either over eighteen or accompanied by their parents", then one would expect to be allowed in when one condition is met, when the other condition is met, and when both conditions are met. That is, the disjunction is naturally taken to be inclusive. On the other hand, in the sentence "A party must either poll more then five percent of the vote, or lose their deposit", the disjunction is naturally taken to be exclusive: one would be surprised if a party both polled more than five percent and yet still lost their deposit.

---

[7] Certainly, the ambiguity of English still needs to be addressed and translating from English to symbolic expressions and vice versa is an important skill, but these can be left until after the student is clear about the definitions of the symbols.

In teaching formal logic it is not difficult to avoid confusion on this point. Once the concepts of inclusive and exclusive disjunction are identified, it is reasonable to explain that while in English the two concepts are both expressed by the word "or", in formal logic the two must be kept separate. We then define the "v" operator to be inclusive[8], and show that if we need an exclusive-or we could either invent a different symbol or explicitly write (PvQ)&~(P&Q). The choice of making the basic disjunction inclusive rather than exclusive can be justified by reference to the set interpretation suggested in Section 3.3.4.7: that is, to see disjunction as the correspondent to set union, which is clearly inclusive.

3.5 Difficulties with the Truth-Functional Definition of Conjunction

Even a truth-functional definition of conjunction is not free of problems. Truth-functional conjunction is commutative (P&Q ≡ Q&P), whereas this is not always the case in English. For instance the proposition "She got pregnant and married" is not equivalent to "She got married and pregnant", and "He took off his shoes and went to bed" is not the same as "He went to bed and took off his shoes."

In such examples, the connective "and" carries the sense of "and then". If we distinguish such a time dependent conjunction from the normal conjunction which is independent of time, then the difficulty is minimised. The truth-functional definition of conjunction does accurately represent the latter.

---

8 The symbol "v" comes from the Latin "vel", which is always inclusive.

Chapter 4: The Theoretical Basis of Possible Models Diagrams

4.1 Preface

Having described various existing approaches to the teaching of propositional logic, we now introduce a new approach. The approach involves the representation of propositional wffs by the construction of simple graphs called Possible Models Diagrams (PMDs). This chapter defines PMDs in a theoretical manner which is not the most useful approach for teaching. The task of presenting PMDs in a form suitable for students in an introductory logic course is left to Chapter 5.

This chapter was originally written as a paper to appear as "Visualising Boolean Operations on a Hypercube" in the journal of Mathematical and Computer Modeling [CLAR94]. Some small changes have been made to adapt the original paper to this dissertation.

4.2 Introduction

Although it is known that every boolean function can be represented as a subgraph of a hypercube [HARA89], the implications of this fact have been little explored. That is, how does one construct such graphs and what does one do with them once they have been constructed? Representing boolean functions as graphs may provide an alternate scheme for automated propositional theorem proving, has been used to define concurrent processing [GUPT93] and has useful pedagogical implications [CLAR93, see Chapter 5].

This paper explains how propositional expressions can be represented by graphs which I call Possible Models Diagrams, and defines procedures for combining such graphs according to the standard boolean operators.

4.3 Any Boolean Function can be Represented by a Graph

The hypercube, or n-cube Q_n is a graph of order 2^n whose vertices are represented by n-tuples $<x_1,x_2,...x_n>$ where $x_i \in \{0,1\}$, and whose edges connect vertices which differ in exactly one term. Figure 4.1 shows Q_3, where each vertex is labeled with an abbreviated triple showing the values of x_1,x_2,x_3 respectively (eg we write the label "101" as an abbreviation for the triple $<1,0,1>$).

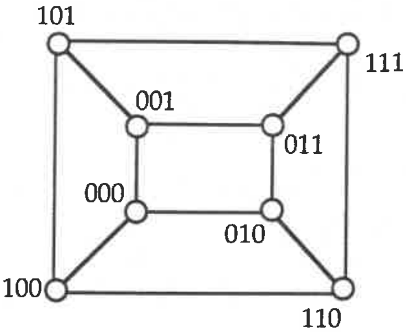


Figure 4.1 — The 3-cube Q_3

For every boolean function there is an equivalent function in the disjunctive form:

$$f(x_1,x_2,...x_n) = \varepsilon_{m-1}x_1x_2...x_n + \varepsilon_{m-2}x_1x_2...x_{n-1}x_n' + \varepsilon_{m-3}x_1x_2...x_{n-1}'x_n' + ... \\ ... + \varepsilon_1 x_1'x_2'...x_{n-1}'x_n + \varepsilon_0 x_1'x_2'...x_n' \quad (4.3.1)$$

where $\varepsilon_i \in \{0,1\}$ for i=1,2,...,m and $m=2^n$ [compare with HARA89].

Harary has shown that if a boolean function of *n* variables is written in this form, then the terms of the disjunction which have $\varepsilon_i=1$ can be used to select a subset S of the vertices of Q_n. The subgraph induced by S can then be seen as a representation of the original boolean function.

For example, the function:

$$f(x_1,x_2,x_3)=x_1x_2x_3 + x_2'x_3 \quad (4.3.2)$$

could be re-written as:

$$f(x_1,x_2,x_3)=1.x_1x_2x_3 + 0.x_1x_2x_3' + 1.x_1x_2'x_3 + 0.x_1'x_2x_3 + 0.x_1x_2'x_3' + 0.x_1'x_2x_3' \\ + 1.x_1'x_2'x_3 + 0.x_1'x_2'x_3' \quad (4.3.3)$$

and hence represented as the induced subgraph of Q_3 whose vertices are 111, 101, and 001 as in Figure 4.2.

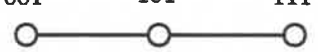


Figure 4.2 — An Induced Subgraph of a Boolean Function

Perhaps we can more clearly illustrate the relationship between Q_3 and the subgraph representing (4.3.3) by leaving those vertices corresponding to terms with $\varepsilon_i=0$ open, while darkening the vertices for terms with $\varepsilon_i=1$, as shown in Figure 4.3.

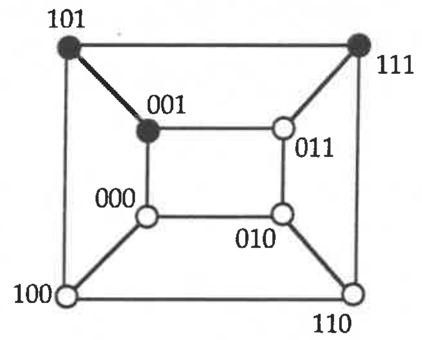


Figure 4.3 — A More Visually Expressive Representation

Given such an induced subgraph, the boolean function can easily be retrieved: simply form a disjunction whose disjuncts are given by the vertices in the subgraph (in any order). If the subgraph is identical to Q_n then the function must be tautologous; if the subgraph is empty then the function must be inconsistent; in other cases the function must be contingent.

However, given an arbitrary boolean function, it would seem from Harary's approach [HARA89] that one must first do substantial algebraic manipulation to find an equivalent function in the form of (4.3.1) before the subset S (and hence the subgraph induced by S) can be determined. This paper will describe another procedure for constructing this induced subgraph, and such a procedure can actually be used as a method for finding an equivalent function in the form of (4.3.1).

4.4 Any Propositional Formula Can be Expressed as a Boolean Function

It is implied by (4.3.1) that all boolean functions can be expressed using only the operators negation, conjunction and disjunction. Nevertheless, there are other boolean operators (such as material implication and the bi-conditional) which play an important part in many systems of logic and many logical formulae are more naturally expressed using such operators. In order to broaden the applicability of hypercubes to propositional logic, we will also want to incorporate these operators.

Definition 4.1 [compare with LEMM65 p44]

A well-formed formula (wff) of propositional logic is defined recursively as —

i) Any propositional variable on its own is a wff.

ii) If α is a wff then so is $\sim\alpha$.

iii) If α and β are wffs then so are $(\alpha \& \beta)$, $(\alpha \vee \beta)$, $(\alpha \Rightarrow \beta)$ and $(\alpha \Leftrightarrow \beta)$.

iv) No expression is a wff except in virtue of i), ii), and iii).

Several notes should be made about this definition:

- Whereas truth values in propositional logic are normally designated as True and False, there is a one-to-one correspondence between these values and the 1 and 0 of boolean algebra.
- In accordance with the usual truth-table definitions of the operators ~ (negation), & (conjunction), v (disjunction), ⇒ (material implication), and ⇔ (bi-conditional), each wff with *n* distinct propositional variables $V_1, V_2, \ldots, V_n$ defines a function from $V_1 x V_2 x \ldots x V_n$ into {0,1}.
- All the logical connectives (apart from negation) are formally defined as dyadic operators. However, the propositional conjunction operator is associative and hence the wff $(V_1\&(V_2\&V_3))$ can be written as the unbracketed product $V_1V_2V_3$ in boolean algebra. Likewise, nested propositional disjunctions may be written as an unbracketed sum in boolean algebra.
- Although this definition employs operators other than negation, disjunction and conjunction, such operators may still be defined as boolean functions in the form (1). For instance the material implication operator ⇒ can be defined as the function —

$$\text{Implies}(x_1,x_2) = x_1' + x_2 = 1.x_1x_2 + 0.x_1x_2' + 1.x_1'x_2 + 1.x_1'x_2' \quad (4.4.1)$$

Given Definition 4.1, any wff containing *n* distinct propositional variables may be expressed equivalently as a boolean function of *n* variables, requiring only negation, disjunction and conjunction.

4.5 Constructing the Graph of a Propositional wff

Since all propositional wffs have an equivalent boolean function, and every boolean function can be expressed in the form (4.3.1), and every function in the form of (4.3.1) can be represented by an induced subgraph of a hypercube, it follows that every propositional wff may be represented as an induced subgraph of a hypercube. In this section I describe two procedures for constructing such a graph: the first is most useful as a visual procedure, while the second is more useful if the task is to be computerised.

4.5.1 Hypercubes are Better Drawn Hierarchically

Throughout the rest of this paper I will use *n* to denote the number of distinct propositional variables in a particular wff, thus allowing the wff to be shown as a subgraph of a hypercube of degree *n*. These hypercubes will be drawn hierarchically rather than in the usual form with no line-crossings: that is, instead of drawing Q_3 as in Figure 4.1, I will draw it as shown in Figure 4.4.

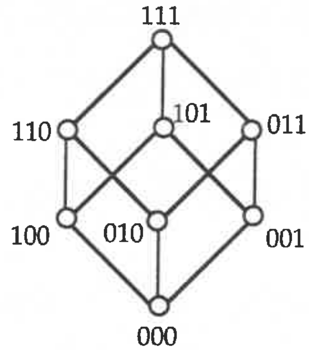


Figure 4.4 — Preferred Diagram for Q_3

In these hierarchically-drawn hypercubes, vertices are arranged into $n+1$ rows ($R_0, R_1, \ldots, R_n$) such that R_i contains those $\binom{n}{i}$ vertices whose n-tuple contains exactly i zeros. Thus, the top row always contains the single n-tuple <1,1,...,1> and the bottom row always contains just <0,0,...,0>. Within any row, the vertices are arranged from left to right in decreasing order of magnitude (when the vertices' labels are considered as binary numbers). The reason for this preference is that it allows hypercubes of any size to be constructed and labeled in a consistent manner, and that it allows labels to be assumed rather than explicitly written.

A hypercube drawn in this manner, with the possibility that some vertices are darkened and others left open, I call a Possible Models Diagram (PMD), for reasons which will later be clear. A PMD is a graph $G \cong Q_n$ (ie a graph isomorphic to a hypercube) whose vertex set V is partitioned into two sets, T(G) (the darkened vertices) and F(G) (the vertices left open).

A PMD may also be interpreted as a Hasse diagram: that is, the hierarchical structure may be seen as a partial ordering of possible models. Brink and Heidema have applied this type of ordering to the verisimilitude of scientific theories [BRIN87].

4.5.2 Each Propositional Operation Corresponds to a Visual Manipulation of Hypercubes

Following Definition 4.1, we can form induced subgraphs of Q_n as follows —

i) Every propositional variable on its own can be represented as a subgraph containing exactly half the vertices of Q_n, namely, those vertices for which that propositional variable is True (ie 1). In other words, a propositional variable V_i can be represented by a PMD X in which $T(X)=\{<V_1,V_2,\ldots,V_n> \mid V_i=1\}$. For example, in a system with two propositional variables, the propositional variable V_1 may be represented by the induced subgraph shown in Figure 4.5a, or more clearly by the PMD in Figure 4.5b.

Figure 4.5a. Induced Subgraph for V_1 in a System of Two Variables

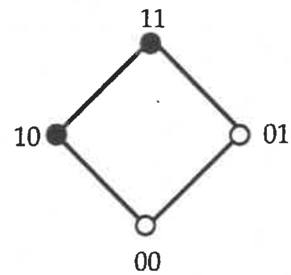


Figure 4.5b. Possible Models Diagram for V_1 in a System of Two Variables

ii) Negation: if the wff α is represented by the PMD A, then the wff ~α can be formed by reversing the open and darkened vertices of A. That is, by constructing the PMD X with T(X)=F(A).

iii) If the wffs α and β are represented by the PMDs A and B respectively, then the PMDs for (αvβ), (α&β), (α⇒β) and (α⇔β) can be constructed using the following visual operations —

Disjunction: form the PMD for (αvβ) by overlaying A onto B. That is, form a PMD X with T(X)=T(A)∪T(B).

Conjunction: form the PMD for (α&β) by matching-the-dots from A and B. That is, form a PMD X with T(X)=T(A)∩T(B).

Material implication: form the PMD for (α⇒β) by reversing A and overlaying onto B. That is, form a PMD X with T(X)=F(A)∪T(B).

Bi-conditional: form the PMD for (α⇔β) by finding vertices in A and B which have equal value. That is, form a PMD X with T(X)=(T(A)∩T(B))∪(F(A)∩F(B)).

4.5.3 Constructing a PMD (Algorithm 1)

Given a wff containing *n* distinct propositional variables, a PMD for that wff may be constructed by hand as follows —

- Draw an appropriate PMD underneath every propositional variable in the wff
- While (there is an operator for which no PMD has yet been constructed but whose operand(s) have PMDs already) do
    - Merge the PMDs of the operand(s) using the appropriate visual operation defined above

  end while
- The graph which was drawn last is the PMD for the wff

For example, consider the wff P⇒(PvQ). Since the wff contains two distinct propositional variables, we can represent it as an induced subgraph of Q_2 and this subgraph can be constructed as a PMD as shown in Figure 4.6.

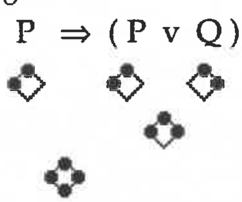


Figure 4.6 — Constructing the PMD for a wff

Note:
- When using this algorithm by hand, the vertices need not be labeled: the labels are assumed to follow the convention described earlier.
- The final PMD in Figure 4.6 has all vertices darkened, implying that the wff P⇒(PvQ) is tautologous.

One can now see the purpose of the nomenclature: a Possible Models Diagram is a graph G which has one vertex for each possible model of *n* variables; and darkened vertices (those belonging to T(G)) indicate all those models in which the wff is True. Even though many wffs may have the same PMD, any two wffs with the same PMD will be truth-functionally equivalent. That is, each possible PMD represents a truth-functional equivalence class of wffs.

4.5.4 Constructing a PMD (Algorithm 2)

If the construction of PMDs were to be computerised, then the following recursive algorithm may be used —

- Parse the wff and build a binary tree with operators as nodes and propositional variables as leaves
- To construct a PMD for the root node —
  - If this node's operator is not Negation then
    - Construct a PMD for the left subtree

    end if
  - Construct a PMD for the right subtree
  - Merge the two subtree PMDs in accordance with the operation defined above for this node's operator

This algorithm is simply a left-to-right depth-first traversal of the binary tree which reflects the structure of the wff.

4.6 A Comparison of Possible Models Diagrams and Truth Tables

There is, of course, a strong similarity between PMDs and Truth Tables. Both indicate the value of a propositional expression for every possible model: whereas a PMD has 2^n vertices, the corresponding truth table has 2^n rows. Either could be used to define the boolean operators. Once a PMD has been constructed, it is a trivial matter to copy the information into a truth table, and vice-versa.

There are, however, two points in the favour of PMDs. One is that the visual nature of PMDs gives some pedagogical advantage over truth tables [CLAR93, see Chapter 5]. The second relates to the way sequents of propositional logic can be analysed by PMDs, as described in the following section.

4.6.1 PMDs can be used to Prove the Validity of Propositional Sequents

If $A_1, A_2, A_3, \ldots, A_m$ and B are propositional wffs, then $A_1, A_2, A_3, \ldots, A_m \therefore B$ is a propositional sequent claiming that the assumptions $A_1, A_2, A_3, \ldots, A_m$ entail the conclusion B. The validity of such a claim may be established by showing that the corresponding conditional $(A_1 \Rightarrow (A_2 \Rightarrow (A_3 \Rightarrow \ldots (A_m \Rightarrow B) \ldots)))$ is a tautology. As we have discussed, a wff may be shown to be tautologous by hand-constructing either a PMD or truth table, but this may be cumbersome for complex sequents.

An alternate way to check the validity of a sequent is to compare the models for which the left-hand side is true with the models for which the right-hand side is true. If the assumptions are to entail the conclusion, then every model which satisfies the assumptions must also satisfy the conclusion. In other words, the models which satisfy the conjunction[1] of the assumptions must be a subset of the models which satisfy the conclusion.

In terms of PMDs, a sequent $A_1,A_2,A_3,\dots,A_m$ ∴ B may be analysed as follows. First construct a PMD L for the wff $A_1\&(A_2\&(A_3\&\dots\&A_m)))$ and a separate PMD R for the conclusion B. The sequent is valid if and only if $T(L)\subseteq T(R)$.

For example, suppose we want to test whether the sequent P, ~(P&Q) ∴ ~Q is valid. First, construct a PMD for the conjunction of the assumptions and a separate diagram for the conclusion (see Figure 4.7).

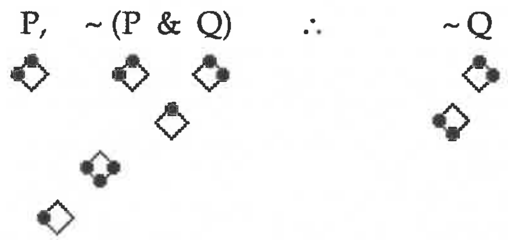


Figure 4.7 — Example of Proving Sequent Validity

In this example, T(L)={<1,0>} and T(R)={<1,0>,<0,0>}. Since $T(L)\subseteq T(R)$, the sequent P, ~(P&Q) ∴ ~Q must be valid.

This form of sequent validation could also be carried out using truth tables, but once again it is hoped that the visual effect of the PMDs makes it somewhat easier. The same thing could also be achieved using boolean functions in this way — find a function in the form of (4.3.1) which is isomorphic to the conjunction of the assumptions; find another function in the form of (4.3.1) which is isomorphic to the conclusion; then the sequent is valid if and only if the first function is subsumed by the second.

4.7 Conclusion

Every propositional wff can be expressed as a boolean function, and hence as an induced subgraph of a hypercube. A Possible Models Diagram is a hypercube in which vertices may be darkened to indicate this induced subgraph. Given any wff, there is a straight-forward procedure for constructing that wff's PMD and this procedure can be executed either by hand or by computer.

PMDs can be used just as readily as truth tables to establish whether a wff is tautologous, contingent or inconsistent. Furthermore, PMDs may be used to prove the validity of any propositional sequent.

---

[1] The commas between assumptions on the left-hand side of the sequent are taken as implicit conjunctions.

Chapter 5: Presenting Possible Models Diagrams to Students

5.1 Preface

This chapter was presented at the ACM's 24th SIGCSE Technical Symposium on Computer Science Education and later published as "Possible Models Diagrams: a visual alternative to truth tables" in the SIGCSE Bulletin, Volume 25, Number 1, March 1993 [CLAR93]. Some small changes have been made to adapt the original paper to this dissertation.

Although the content of this paper covers some of the same ground as the previous chapter, it presents the ideas in a form which is more appropriate to teaching than the previous technical approach. Whereas the previous Chapter drew a parallel between PMDs and boolean functions, this Chapter emphasises the parallel between PMDs and set theory.

5.2 Introduction

It is standard practice to introduce university students to truth tables quite early in computer science courses. Truth tables provide a convenient way to teach boolean logic, which forms the basis of the digital computer. Truth tables also introduce propositional logic concepts and are often presented along with some formal proof structure and associated rules of derivation.

Later, the limitations of propositional calculus lead to the need for predicate calculus and here truth tables must be put aside. Nevertheless, the concepts described by truth tables form essential groundwork for any career in either logic or computing.

Regardless of the fact that truth tables are virtually universally standard, they aren't so sacrosanct as to make alternatives unthinkable. Indeed truth tables do present some teaching problems which allow for some improvement. In particular, the truth table definition of material implication (and disjunction to a lesser extent) always confuses students. The fact that an implication should be considered true when the antecedent is false is neither intuitive nor easily remembered. (See Chapter 3.) In the interest of improved teaching, this paper presents an alternative which is at least as expressive as truth tables, yet more intuitive to the novice and more easily remembered by the student.

Section 5.3 presents a student's-eye-view of Possible Models Diagrams (PMD), and then Section 5.4 supports the logical soundness of the approach.

5.3 Possible Models Diagrams: a Student Guide

Suppose P and Q represent two propositions. Then there are four possible states of the world: the state in which both P and Q are true; the state where P is true but Q is false; the state where P is false but Q true; and the state where both P and Q are false. We could show this situation in a simple graph shown in Figure 5.1.

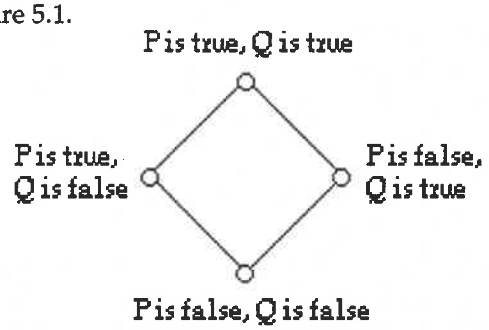


Figure 5.1 — The Four Possible Models of P and Q

Each of these possible situations is called a model (or an interpretation) of P and Q and hence the graph is called a Possible Models Diagram (PMD). We show that some expression of propositional logic is true for a particular model by filling in the corresponding node of the graph. Thus the simple propositions P and Q are represented by the PMDs shown in Figure 5.2a and Figure 5.2b respectively.

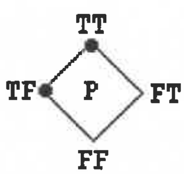


Figure 5.2a — PMD for P

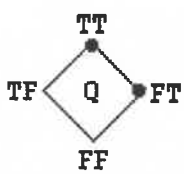


Figure 5.2b — PMD for Q

5.3.1 Combining Rules

We have seen that Possible Models Diagrams are an easy way of visualising simple propositions and now show a number of ways of combining these diagrams to represent compound statements.

5.3.1.1 Overlaying two diagrams has the same effect as disjunction (Figure 5.3). Note the similarity between overlaying and the set union operation.

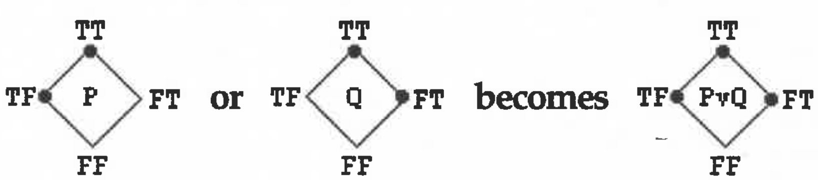


Figure 5.3 — Overlay — the Visual Operation for Disjunction

5.3.1.2 Finding the corners which match on two diagrams has the same effect as conjunction (Figure 5.4). Note the similarity with the set intersection operation.

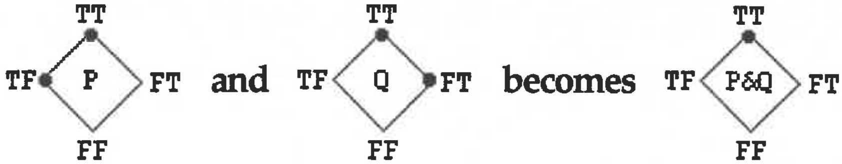


Figure 5.4 — Match — the Visual Operation for Conjunction

5.3.1.3 Reversing each corner of a diagram is the same as negation (Figure 5.5). Note the similarity with set complementation.

Figure 5.5 — Reverse — the Visual Operation for Negation

5.3.1.4 How do we capture the idea of a subset in propositional logic? Suppose that some set P is a subset of another set Q: we use the notation $P \subseteq Q$ to mean that "every element of P is also an element of Q". Now suppose we examine various possible entities in the universe and see whether or not they confirm such a conditional claim. We find three situations which are consistent with the subset claim (x_1 such that $x_1 \in P$ and $x_1 \in Q$, x_2 such that $x_2 \notin P$ and $x_2 \in Q$, and x_3 such that $x_3 \notin P$ and $x_3 \notin Q$) and one situation which is inconsistent with the subset claim (x_4 such that $x_4 \in P$ and $x_4 \notin Q$). The entity x_4 could not be positioned anywhere on a Venn Diagram showing $P \subseteq Q$.

These four situations may be related to the four possible models in a PMD: x_1 is the counterpart to the top node of the PMD (in which P is true and Q is also true), x_2 is the counterpart to the right node of the PMD (in which P is false but Q is true), x_3 is the counterpart to the bottom node of the PMD (in which P is false and Q is also false), and x_4 is the counterpart to the left node of the PMD (in which P is true but Q is false). The first three nodes should be darkened to indicate that the material implication $P \Rightarrow Q$ is supported by those models, but the fourth should be left open to indicate that material implication cannot be true in that model[1]. This is shown in Figure 5.6.

[1] It has been suggested to me that I am attempting to define an equivalence between sets and propositional logic which is logically mistaken. A collection of entities forming a *set* P cannot be equated with a *proposition* P. A proposition is a claim that some atomic assertion is true, whereas set membership is a predicate applying to a number of entities. A claim that $x \in P$ is more accurately translated into the logical statement P(x) — a statement of predicate logic rather than propositional logic. This criticism is seen to apply especially to conditional statements. The set expression $P \subseteq Q$ claims a certain relationship between categories ("All P's are Q's"), which, in modern logic, must be rendered as the predicate expression $(\forall x)P(x) \Rightarrow Q(x)$, not as the propositional expression $P \Rightarrow Q$.

I agree that as a technical point of logic, such an equivalence would be erroneous, but such an equivalence is <u>not</u> what I am claiming. Rather, I am attempting to take the students' pre-existing

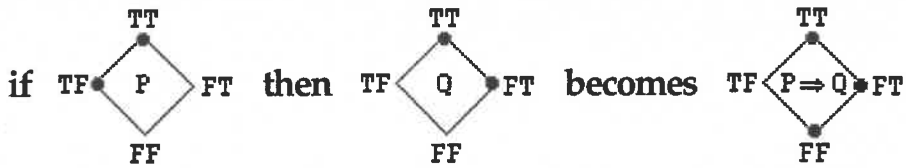


Figure 5.6 — Reverse-and-overlay — the Visual Operation for Material Implication

Two things to note about the operation in Figure 5.6 —

i) It is easy to remember that material implication has the diagram  because the three dots resemble an arrow pointing to the right.

ii) When combining two diagrams as above, the visual procedure is to reverse the first diagram and overlay it on the second[2].

5.3.2 More Complex Expressions

Set equivalence P≡Q is reflected in propositional logic by the bi-conditional: P⇔Q is defined as (P⇒Q)&(Q⇒P). This can be diagrammed by appropriate combinations of the previous operations, as shown in Figure 5.7.

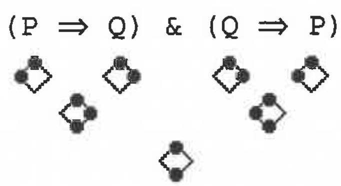


Figure 5.7 — Visual Operations for the Bi-conditional

The final outcome is intuitively sensible, since it shows those models in which the truth of P is exactly the same as the truth of Q.

---

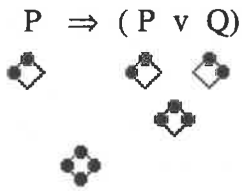
notions of set operations as the foundation on which to build similar concepts in logic. An informal account of set theory would equate the symbolic expression P⊆Q with descriptions such as "the set P is completely inside the set Q" and "anything in the set P must also be in the set Q" and "membership of P implies membership of Q". Thus, the notion of a subset follows the same reasoning schema as the notion of material implication.

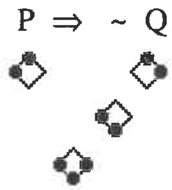
Rather than claiming that set expressions can be expressed equivalently as propositional expressions, I am claiming that set operators invoke the same reasoning schema as propositional operators.

Perhaps an alternate way to dismiss this criticism is to take an approach suggested in Section 2.3.6.2. If P represents some proposition, then there is a natural correspondence between the truth of P and membership of the set of all possible worlds in which the proposition is true.

[2] To the student, the procedure "reverse and overlay" is an easy habit to learn and it is only later that they discover that P⇒Q is logically equivalent to ~PvQ. Some teachers may prefer to use this equivalence as a <u>definition</u> of material implication, in the same way that [COPI61] uses ~(P&~Q) as a definition. I find that students rebel against this use of fiat, but are more ready to accept the parallel between material implication and subset.

In fact a PMD can be constructed using the three basic combining rules (reverse, overlay and match) for any well-formed formulae which uses only two propositional variables. Figures 5.8a, 5.8b and 5.8c show three examples.





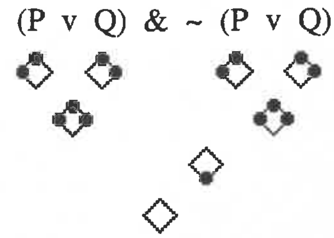


Fig 5.8a — PMD for P⇒(PvQ) Fig 5.8b — PMD for P⇒~Q Fig 5.8c — PMD for (PvQ)&~(PvQ)

Every wff falls into one of three categories —

i) Tautology — a wff whose PMD has every corner filled in (eg Figure 5.8a). ie a wff which is true however you interpret it.

ii) Contingent — a wff whose PMD has some corners filled in but not all (eg Figure 5.8b). ie a wff which is sometimes true and sometimes false.

iii) Inconsistent — a wff whose PMD has no corner filled in (eg Figure 5.8c). ie a wff which is always false.

5.3.3 Generalising to Three Propositional Variables

If there are three propositional variables (say P,Q,R) in a wff, there will be eight possible models. These are best visualised as the corners of a cube whose opposing faces represent P and ~P, Q and ~Q, R and ~R respectively[3]. However, this can be shown diagrammatically as a two-dimensional graph (Figure 5.9).

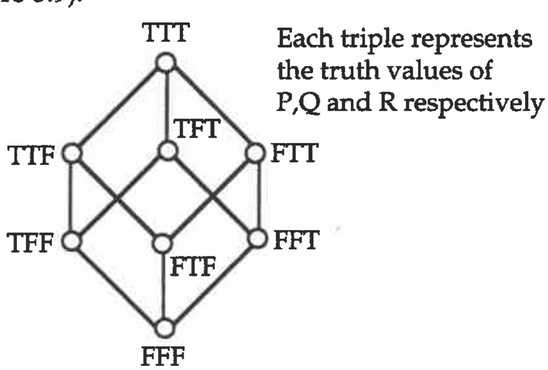


Figure 5.9 — A PMD for Three Variables

The combining rules (reverse, overlay and match) can all be used as before.

[3] In teaching this section a real cube is a useful visual aid (see section 6.4.2).

5.3.4 Provable Sequents

After showing how complex statements may be represented as symbolic propositional formulae, a course in logic is then likely to proceed to the concept of a structured argument. We first describe a sequent in some form like $A_1,A_2,A_3,\ldots,A_n \therefore B$ (meaning "the premises $A_1,A_2,A_3,\ldots,A_n$ entail the conclusion B"). We then describe a range of derivation rules which allow one to proceed logically from one statement to the next in a proof of the sequent.

In any propositional calculus, a sequent $A_1,A_2,A_3,\ldots,A_n \therefore B$ may be shown to be valid (and hence provable, according to the Completeness Theorem) by showing that $(A_1 \Rightarrow (A_2 \Rightarrow (A_3 \Rightarrow \ldots (A_n \Rightarrow B)\ldots)))$ is a tautology, however, this is rather cumbersome in general. PMDs provide an alternate way to check whether a sequent is valid.

Suppose we want to test whether the sequent P, ~(P&Q) ∴ ~Q is valid. First, construct a PMD for the wffs on the left-hand side (the comma is treated as an implicit conjunction) and a separate diagram for the right-hand side (Figure 5.10).

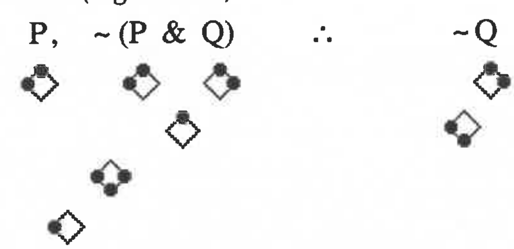


Figure 5.10 — Validation of the Sequent P, ~(P&Q) ∴ ~Q by PMD

Now apply this simple rule: a sequent is valid if and only if the diagram for the premises is a subset of the diagram for the conclusion.[4] In the example above, the possible models represented by  are a subset of those represented by  and hence the sequent is valid.

The reader may like to compare this method of validating the sequent P, ~(P&Q) ∴ ~Q with those described in Section 2.3.

5.4 Possible Models Diagrams: Logical Soundness

5.4.1 Possible Models Diagrams are Really Sets

In order to see that PMDs are a sound way of representing wffs, one should first ignore the lines connecting the graph and simply think of the set of nodes as the set of all possible models of n

[4] At least, that is a simplified statement of the rule which is easily remembered by students. The fully stated rule is "A sequent is valid if and only if the set of models indicated by the PMD for the conjunction of the premises is a subset of the models indicated by the PMD for the conclusion."

propositions, $U_n = \{<x_1,x_2,...x_n> \mid x_i \in \{T,F\} \text{ for } i=1,2,...,n\}$. The lines link elements of this set in a certain way, but this is only important for the visual effect.

For n=2, we get the set of all possible models of two propositions $U_2 = \{<T,T>,<T,F>,<F,T>,<F,F>\}$. Now we need only to equate the first proposition with the set $\{<x_1,x_2> \in U_2 \mid x_1=T\}$ and the second proposition with the set $\{<x_1,x_2> \in U_2 \mid x_2=T\}$.

Suppose W is the set representing some wff. Then the reverse operation is simply $\sim W = \{<x_1,x_2> \in U_2 \mid <x_1,x_2> \notin W\}$. If V is a set representing some other wff, then overlay corresponds to $V \vee W = \{<x_1,x_2> \in U_2 \mid <x_1,x_2> \in V \text{ or } <x_1,x_2> \in W\}$ and match corresponds to $V\&W = \{<x_1,x_2> \in U_2 \mid <x_1,x_2> \in V \text{ and } <x_1,x_2> \in W\}$.

Thus, the operations defined on PMDs correspond to exactly those primitive logical operations we expect.

5.4.2 Converting Between Truth Tables and Possible Models Diagrams

Whereas a PMD represents all possible models as nodes on a graph, a truth table represents them as lines in a table. Given a PMD, it is simple to construct an equivalent truth table: for each node in the diagram, if the node is filled-in, place a T in the corresponding row of the table, otherwise place an F in the corresponding row of the truth table. Converting in the opposite direction is equally trivial. An example is shown in Figure 5.11.

P Q	P & Q
T T	T
T F	F
F T	F
F F	F

Figure 5.11 — Translating Between a PMD and a Truth Table

5.4.3 Generalising to More than Two Propositional Variables

Given that a PMD is just a way of visualising the set of possible models (Section 5.4.1), it should be clear that such diagrams can be formed for wffs containing any number of propositional variables. In general, if a wff contains n distinct propositional variables, then its PMD will have 2^n nodes (just as the corresponding truth table will have 2^n rows).

In teaching situations it is rare to set a truth table problem with more than three or possibly four variables, and the same would apply to PMDs. Both methods generalise in theory, but in practice we would rarely use either method for cases where the number of variables is larger than four.

5.4.4 The Rule for Sequent Validity

In Section 5.3.4 we saw the rule "a sequent is valid if and only if the diagram for the premises is a subset of the diagram for the conclusion." This can be justified by again thinking in terms of sets (however, the justification in Section 4.6.1 is probably simpler).

Suppose the conjunction of premises forms the wff X, represented by a PMD whose nodes form the set $X' \subseteq U_n$, and suppose the conclusion Y gives rise to a Diagram whose nodes form the set $Y' \subseteq U_n$. If $X' \subseteq Y'$ then every possible model of the premises is also a model of the conclusion.

Now if every possible model for which some condition X is true is also a model for which Y is true, then X is a sufficient condition for Y. Hence $X \Rightarrow Y$ is necessarily true in all models (ie it is a tautology). Further, whenever $X \Rightarrow Y$ is a tautology, X ∴ Y is provable (by the Completeness and Deduction Theorems) and so from the fact that one PMD is a subset of a second, we can deduce that the wff represented by the second is derivable from the wff represented by the first.

Conversely, if Y can be derived from X (ie X ∴ Y is a valid sequent), then it must be that $X \Rightarrow Y$ is a tautology. This is the case exactly when ~XvY is also a tautology. Now according to our set interpretation (Section 5.4.1), this is the same as saying that the set $\{<x_1,x_2,\ldots x_n> \in U_n \mid <x_1,x_2,\ldots x_n> \notin X' \text{ or } <x_1,x_2,\ldots x_n> \in Y'\}$ is equivalent to U_n. But this equivalence is only possible when $X' \subseteq Y'$.

5.5 Conclusion

Although truth tables are the standard way of introducing boolean logic, a more visual teaching tool has some advantages. This paper has presented the Possible Models Diagrams as such a visual tool. PMDs are simple to learn and manipulate and more hence more enjoyable and memorable for students. (Chapter 7 seeks to substantiate these claims.)

None of the theoretical basis is lost by using Possible Models Diagrams rather than truth tables: they can both be used to define basic logical operations and to test whether a wff is tautologous, contingent or inconsistent. Furthermore, Possible Models Diagrams can also be used for a task which is difficult for truth tables, namely to checking the validity of a sequent.

Chapter 6: Positioning Possible Models Diagrams in a Computer Science Logic Course

6.1 Introduction

The two preceding chapters build a foundation for PMDs in graph theory, boolean algebra and set theory, and indicate how this approach may be presented to students. This chapter seeks to deal more fully with the question of how PMDs may be integrated into a complete course on introductory logic. This is an essential prelude to Chapter 7 since any evaluation of the PMD approach must be seen in a broader context. This broader context should include the knowledge which students are expected to have before encountering PMDs, the methods of teaching used, and the ways in which ability to use PMDs is applied in later topics.

As discussed in Chapter 1, the PMD approach to teaching propositional logic was developed as part of a first year university course on logic for computer science students in which there was an increasing number of educationally disadvantaged students. This logic course is now described in more detail.

6.2 Outline of Course Content

At the University of Natal, Pietermaritzburg, the Computer Science 1 course is comprised of four sections: Pascal programming, computer systems and architecture, software packages on personal computers, and introduction to logic. The introduction to logic component is 25% of the complete Computer Science 1 course in terms of duration and assessment. It is presented in the second semester and consists of the following topics[1] (in the order of presentation):

6.2.1 Basic concepts

The purpose of logic is discussed, and the basic terminology of "propositions", "arguments", "truth", "validity" and "soundness" is introduced. Common forms of fallacious reasoning are described and examples presented. (8% of course time)

6.2.2 Inductive logic

Principles and examples are discussed for each of the main types of inductive reasoning: generalisation, causality (including Mills' Methods), hypothesis formation and refutation (including an explicit description of abduction), and reasoning by analogy. (17% of course time)

---

[1] For more detailed notes about the content of this module, see the Appendix.

6.2.3 Axiomatic systems

A brief introduction to systems in which mechanical manipulation of symbols allows "theorems" to be derived from "axioms". (2% of course time)

6.2.4 Set theory

An introduction to set terminology (membership, union, intersection, subset, cardinality etc) and symbolism. Venn Diagrams are used to prove important set identities. Some logical paradoxes relating to sets are examined (primarily Russell's Paradox) since these provide good examples of important forms of reasoning and also encourage students to think critically about formal systems. (8% of course time)

6.2.5 Propositional Logic

The propositional connectives are defined in terms of their set counterparts. The representation of well-formed formulae as PMDs and the use of PMDs to categorise well-formed formulae as either tautologous, contingent or inconsistent is demonstrated. Truth Tables are shown to be another way to represent the same information as recorded in a PMD. The usefulness of normal forms for computational logic is discussed. Arguments expressed in English are converted to symbolic form and proved using a natural deduction system. Various meta-theorems about the natural deduction system are proved and the issues of consistency and completeness are briefly described. (33% of course time)

6.2.6 Predicate Logic

Predicates and quantifiers are introduced and the natural deduction system is extended to accommodate them. The use of "interpretations" is discussed in the context of disproving predicate sequents, and again the issues of consistency and completeness are briefly described. The introduction of an identity predicate is also briefly discussed. (17% of course time)

6.2.6.1 Variations to Lemmon

In both the section on Propositional Logic and on Predicate Logic I have employed the natural deduction system of Lemmon [LEMM65]. Lemmon's fourteen rules of derivation are used as originally formulated, but I use a slightly different notation.

Whereas Lemmon allows brackets to be dropped when there is not ambiguity, I require all brackets to be written, so that &,v,⇒ and ⇔ are all explicitly dyadic. Where Lemmon would allow A&(BvCvD) as an abbreviation for either A&(Bv(CvD)) or A&((BvC)vD), I require that such formulae be written in full. This is an relatively unimportant modification, but if students are later required to write computer programs which parse boolean expressions, then the notation will reflect the structure of the resulting binary parse tree.

Secondly, whereas Lemmon uses the notation (x) to indicate that the variable x is universally quantified, I use the notation ($\forall$x). This introduces a redundant symbol, but that symbol is common in other systems [eg HAMI78, ROBI79, MCCA81, PAUL87, LU89 and GRIE93] and forms a nice symmetry with the existential quantifier ($\exists$x).

Thirdly, whereas Lemmon denotes predicates by an upper-case letter followed by a series of variables written in lower case (eg Fxyz), I make greater use of parentheses and commas (eg F(x,y,z)). This makes the notation similar to most computer programming languages, where the predicate is implemented as a separate module of code (a function) and the variables are parameters being passed to that function. Separating the symbols by parentheses and commas allows both the function name and parameter names to be longer than one letter without causing confusion. Thus, one could write the three-place predicate "between(Durban, Pietermaritzburg, Johannesburg)", where Lemmon would be forced to define single-letter names and then write something like "Bxyz". As a consequence of this change in notation, predicate wffs in my system are often longer than the equivalent in Lemmon's, and more care needs to be taken to match parentheses. However, for computer science students this is quite appropriate.

While I have strictly adhered to Lemmon's fourteen rules of derivation, I am impressed by the modifications suggested by Allen and Hand [ALLE92 ppxi–xiv] and would adopt them in future.

6.2.7 Other Forms of Logic

Modal, multi-valued, probabilistic (including Bayes Theorem and Certainty Factors) and fuzzy logics all receive brief treatments. (17% of course time)

6.3 Relevance of Logic to Computer Science

This course is a much broader introduction to logic than is typical in computer science syllabi. Nevertheless, every effort is made to relate each section of the course to applications in computing. Underlying the whole course is the notion that a computer is a machine which rigidly adheres to a highly structured form of logic and that understanding "rigid adherence to logical structures" is an important skill if one is to understand computing. More specifically, many sections of this course lay the groundwork for later topics in computing, as summarised in Table 6.1. In addition, examples throughout the course frequently relate to computing topics. However, one can see from responses to the Course Evaluation Questionnaire (see Tables 7.12 and 7.13) that some students continue to feel that the course is not directly relevant to their aspirations as computer scientists.

Table 6.1 — Areas in Computing which Follow-on from the Logic Course

Topics covered in the logic course	Some areas in computing which apply these topics
Causation	Locating and rectifying computer bugs
Hypothesis formation and testing	Locating and rectifying computer bugs
Analogical thinking	Case-Based Reasoning
Abduction	Techniques in Knowledge-Based Systems for classification and diagnosis
Axiomatic systems	Computers process patterns of high and low voltage which are inherently meaningless but imputed with meaning by humans
Set theory	Various programming languages data structures and operations for sets (eg Pascal, Lisp) Grammars, automata and theory of computability Paradoxes of set theory relate to meta-theory regarding the limits of formal systems
Two-valued logic	Boolean data types Digital logic and circuit design
Truth tables and PMDs	Boolean algebra, gate networks and circuit design Boolean expressions in high-level programming languages
Normal forms of logical expressions	Resolution Automated theorem proving Programming in Prolog
Propositional and predicate logic as a whole, and the concepts of consistency and completeness	Formal program verification – Program specification in Z – Assertions – Pre- and post-conditions – Calculus of weakest pre-conditions
Predicates	Boolean functions with parameter passing
Translating English into symbolic notation	Writing detailed specifications, pseudo-code and programs from English requirements specifications
Natural deduction system	Ability to follow rigid syntax, which is nearly always required by computing systems
Sequent Introduction with substitution	The use of sub-programs with parameters
Free and bound variables	Constants and variables in third-generation languages Instantiation of variables in Prolog
Modal, multi-valued, probabilistic and fuzzy logics	Techniques in Knowledge-Based Systems and Artificial Intelligence Reasoning under uncertainty

6.4 Teaching Methods

6.4.1 Lectures and Tutorials

At the beginning of this course students are given 71 pages of notes, which are reproduced in the Appendix. The class is told that these notes are to be the main source of information for the course, rather than the lectures. Each week some section of these notes are designated as required reading. At the beginning of the first lecture of the week, the whole class is required to answer a brief quiz consisting of three or four short answer questions and up to eight true/false questions. These are marked immediately and the answers are discussed during the lecture period. The marks for these quizzes are recorded and constitute a small percentage of the students' final grades.

After the quiz, the class is asked to indicate which sections of the reading were not completely clear and these topics are listed on the board. The remaining time in that lecture period and in the following lecture period is then allocated to clarifying and expanding on those topics. A copy of the course notes on overhead transparencies is often used to reinforce and expound on what the class has already read. Examples selected from the course notes are discussed in finer detail and further examples are invented by both the lecturer and the class.

During the weekly tutorial period, the class primarily works on exercises taken from the course notes, though there will also be further discussions and explanations by the lecturer. The difference between a "lecture" and a "tutorial" is minimised in this course, so that a "lecture" will certainly include discussion and exercises, and a "tutorial" may contain some lecturing.

6.4.2 Teaching Aids

We also make use of a computer program called LemmonAid[2] during the sections on propositional and predicate logic. LemmonAid is a proof validator for the natural deduction system described in [LEMM65]. This program is introduced in the tutorial periods and the class is given a variety of exercises to be completed in their own time. Student responses to this software have been very favorable (see Section 7.5).

A 30cm transparent perspex cube was especially constructed to demonstrate the idea of a PMD for three variables (Figure 6.2). This cube has black tape around the edges and red labels P,~P,Q,~Q,R,~Q on respective faces. Black markers made from table-tennis balls can be attached to any vertex. When held at the correct angle, the class can see how this cube matches the diagram in Figure 5.9. A propositional wff can be written on the board and then represented on

---

2 LemmonAid was written by John Slaney <John.Slaney@arp.anu.edu.au> at the Australian National University in Canberra. He is quite willing to make the software available to other universities.

the cube by attaching markers to appropriate vertices. Conversely, the lecturer can attach markers in some arrangement and then ask the class to write down a wff which expresses that arrangement.

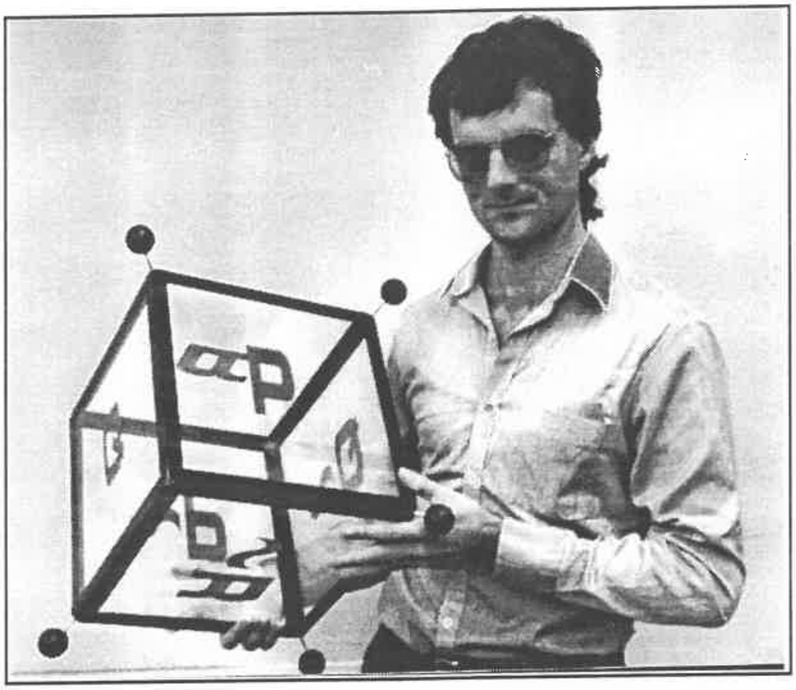


Figure 6.2 — The Author Holding a Perspex Model of a PMD for Three Variables

6.4.3 Assessment Techniques

Students are assessed in four different ways:

- One half of the final three-hour exam;
- Two 45 minute tests;
- A written assignment; and
- The average mark for the weekly quizzes.

In accordance with university policy, the final exam constitutes the major proportion (two thirds) of their final grade, although I would prefer to place more weight on the other forms of assessment.

At the end of the course, the class is given a Course Assessment Questionnaire which enables them to comment on both the course content and the teaching style. This Questionnaire is discussed in Section 7.5.

6.5 Educational Principles Underlying this Course

6.5.1 Interaction in a Social Context

The teaching style of this course has moved away from the traditional lecture style in which the students' role is simply to listen. "Lectures" are interactive periods in which there is opportunity for students to ask questions, times when students learn by working through examples, questions asked by the lecturer in order to gauge the class's understanding, along with some instruction by the lecturer. The class is empowered to choose which topics need further attention, although the syllabus is still ultimately controlled by the lecturer.

The interactive style stresses the active role of the student in learning and, from the outset, the class is informed that learning is primarily their responsibility rather than the teacher's.

It is the view of Vygotksy [VYGO35] and his followers that learning is a socially mediated processes. The use of an interactive style which makes frequent use of question-asking, emphasises the role of the classroom as a social context in which learning may occur.

Vygotsky used the phrase "the zone of proximal development" to indicate "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" [VYGO35, p86]. In other words, tasks in the zone of proximal development are those which are currently beyond the student's ability, but which the student could perform with assistance from a teacher. "The zone of proximal development defines those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are currently in embryonic state" [VYGO35, p86].

Interaction between students and teacher (especially question-asking by the teacher) allows the teacher to monitor and respond to the students' zone of proximal development. In order to enhance learning, the teacher should first establish what is already known by the students, and what the students could perform or learn independently. The teacher can then pitch further instruction at a level just beyond the students' unaided capabilities. (A more detailed analysis of the role of question-asking is given in [GAVE85].) In traditional, non-interactive lecturing, it is inevitable that the instruction is pitched beneath the zone of proximal development (in which case the students become bored) or beyond it (in which case the students cannot absorb the material being presented).

This form of interactive teaching is most suited to small groups of students, or to groups in which all students are progressing at a similar pace. It is made more difficult by having a large group (the courses described here typically have a class-size of about 50), and by having students of diverse abilities and backgrounds (the classes at the University of Natal have a mixture of races, fluency in English, and educational background). Nevertheless, it has still been possible to engage such large, diverse groups and to instill an atmosphere of co-operation, enthusiasm and enjoyment.

6.5.2 Multiple Learning Modalities

In designing an introductory logic course, Utzinger writes:

> The material that is presented should reach the student in as *many modes* as possible. ie he should hear it, see it, write it, say it, feel it and if possible smell it, preferably at the same time. [UTZI82 p10]

Although Utzinger's logic course was primarily designed for learning-disabled students, he is sure that such principles "would also benefit the 'normal' student as well" [UTZI82 p9].

Wittrock points out that "a variety of teaching methods will be needed by each student, depending on his background and its relationship to the subject matter" [WITT74 p182].

In this logic course, students may learn through reading (the course notes, other available textbooks, overhead transparencies, worked examples on the board), listening (to the lecturer and to other students), watching (the PMD cube described in Section 6.4.2), working on examples (in class time as well as in their own time, on paper as well as with the LemmonAid software), and individual discussions with the lecturer.

When working through examples during class time, problem solving strategies are discussed just as much as technical details. In this way it is hoped that students can learn in an apprenticeship style. In part this means that by hearing the lecturer verbalise his strategies (including mistakes and ideas which lead to dead-ends), they will not only learn the right answer, but also learn a way of thinking about logical problems. A good description of the principles and intentions of "cognitive apprenticeship" is given in [COLL85].

6.5.3 Encouraging Transfer of Knowledge

Utzinger claims that translating between English and symbolic notation is much harder for students than learning to understand basic logical principles [UTZI82]. That is, learning the symbolic techniques is easier than applying those techniques appropriately. This comment emphasises the importance of transfer. That is, the importance of students being able to take the principles taught in a logic course and transfer them to other situations. If students can solve problems within some symbolic system but cannot (or do not) make use of the same principles in their everyday experience, then it is unclear what exactly they have learnt, and it is doubtful that the teaching of the symbolic system has been of any real use. Various authors have commented on the low amount of transfer of logical principles [eg GROS90 and NISB87].

The choice of Lemmon's natural deduction system is primarily based on an intuition that a formal method which reflects natural human reasoning schema will promote transfer more readily than a method which differs from those schema. [NISB87] supports such an intuition. Methods such as resolution and semantic tableaux may be easier to teach, but they are quite remote from normal human reasoning (see Section 2.3). That these methods are easy to automate by computer emphasises the fact that they are mechanical procedures. Mechanical procedures are inherently specific to one system and not transferable to other notations or domains. In contrast, a natural deduction system requires much thought and creativity, and hopefully teaches the student about underlying principles rather than about mechanisms only.

Both truth tables and PMDs are essentially mechanical and in teaching them there is always the risk that students will learn the mechanisms but not the underlying concepts. To avoid this it is crucial to emphasise to students that they must interpret what they are doing and translate the results into English, rather than just following the procedures blindly.

It has not been the task of this research to evaluate the degree to which knowledge of formal logic is transferred to other situations. However, Section 8.3 suggests this as an important avenue for future research.

6.5.4 Appropriate Choice of Knowledge Representation

The underlying problem-solving strategy presented throughout this logic course follows three basic steps —

1. Translate the task from English into some symbolic notation
2. Solve the symbolic task using well-defined symbolic operations
3. Translate the solution back into English

This strategy has wide applicability, but for any particular task an appropriate system of symbolic representation and manipulation must be chosen. Knowledge can be represented in a variety of forms and both computer scientists [eg BROW75] and educationalists [eg BRUN67 and LARK87] are aware of the importance of choosing an appropriate knowledge representation. Indeed, the search for appropriate representations is one of the central issues in computing. An appropriate knowledge representation makes learning and retention easier, simplifies problem solving, and enhances the ability to transfer the knowledge to other situations. However, it is rarely the case that one representation of some piece of knowledge is always the most appropriate. Rather, different representations suit different purposes.

For this reason, there are a variety of logic notations and a variety of proof techniques, just as there are a variety of programming languages. In Chapter 2 we noted a large variety of approaches to teaching logic, each with its own advantages and disadvantages. The course under discussion emphasises that there is no one supreme form of logic, but rather that different forms of logic (inductive and deductive, propositional and predicate, as well as modal, probabilistic, multi-valued and fuzzy logics) suit different purposes.

In order to discuss the relative merits of different representations, we will examine two ways of characterising representations: one given by Larkin and Simon in [LARK87] and the other by Bruner in [BRUN67].

First, the concepts of informational and computational equivalence are suggested by Herbert Simon —

> Two representations are informationally equivalent if all the information in the one is also inferable from the other, and vice versa. ... Two representations are computationally equivalent if they are informationally equivalent and, in addition, any inference that can be drawn easily and quickly from the information given explicitly in the one can also be drawn easily and quickly from the information given explicitly in the other, and vice versa. [LARK87 p67]

With these definitions one can see that many of the techniques described in Section 2.3 are informationally equivalent to each other, but typically not computationally equivalent. PMDs and truth tables are informationally equivalent, but the suggestion of this thesis is that they are not computationally equivalent.

In [LARK87], Larkin and Simon present a computational account of problem solving using sentential descriptions of a problem on the one hand and diagrammatic representations on the other. The term "sentential" does not imply "English sentences" but rather a sequence of symbolic expressions derived from the English description of a problem. In the examples they discuss, the two forms of representation are informationally equivalent and the main point of the paper is to evaluate the relative computational power of the two representations. They split problem solving into the tasks of search, recognition and inference and conclude that a diagrammatic representation reduces the need for search and improves the efficiency of recognition, but has little effect on inference processes.

Larkin and Simon claim that an appropriately constructed diagram requires less labeling than an informationally equivalent sentential representation. If the labeling is implicit in the structure of the diagram, then it can still be retrieved when required, but need not require cognitive attention otherwise. Further, they claim that a perceptual enhancement arises from the use of a diagram: that is, a problem solver can perceive important information about a problem more readily from a diagram than from a sentential representation.

The point has been made before that labels may be omitted from PMDs without loss of information. The necessary operations of reverse, match and overlay can be carried out without requiring the confusion of labels, but the labels can be inferred whenever necessary. By contrast, labels cannot be omitted from a truth table: even when the arrangement of rows follows some convention, the processing of connectives requires continue referral to the labels.

(It should be noted that research by Markovits [MARK86] among school children indicates the opposite of Larkin and Simon's thesis. Markovits presented one group of subjects with verbal instructions about a conditional reasoning task, and another group with both verbal instructions as well as pertinent hand-drawn pictures. He found that the second group's performance was markedly *inferior* to the first group's. I think this study can be discounted on the grounds that the drawings were too specific (they were simple drawings of cats etc) and made the subjects think of just one specific situation instead of being general diagrams which may have encouraged the abstraction of underlying principles.)

Second, Jerome Bruner proposes that knowledge representations are characterised by their mode, their economy and their effective power [BRUN67 p44]. The mode of a representation is either enactive (ie a set of actions), iconic (ie a set of summary images that stand for concepts without fully defining them), or symbolic (the use of arbitrarily assigned symbols along with rules for manipulating those symbols) [BRUN67 pp10–11]. In defining economy, Bruner writes "The more items of information one must carry to understand something or deal with a problem, the more successive steps one must take in processing that information to achieve a conclusion, and the less the economy" [BRUN67 p45]. The effective power of a representation is the degree to which the representation's theoretic capabilities are actually put to work by the learner. "The power of a representation can also be described as its capacity, in the hands of a learner, to connect matters that, on the surface, seem quite separate" [BRUN67 p48] — that is, the transferability of the knowledge. It may seem that effective power is a function of the learner rather than a characteristic of the representation, but in reality it is dependent on both. Some students will be able to transfer knowledge more effectively than other students, and some forms of representation will enhance the process of transfer more than others.

Regarding mode, PMDs are primarily iconic: the visual patterns of the basic PMDs for P, Q, ~P, PvQ, P&Q, and P⇒Q are iconic representations of the underlying concepts. In addition, PMDs are partly enactive since the operations reverse, overlay and match are procedural actions which manipulate the icons. In contrast, truth tables are purely symbolic.

Regarding economy, it was earlier claimed that the information required to establish whether or not a sequent is valid is more readily extracted from a set of PMDs than from a truth table because visual information is more quickly processed than symbolic information. To use Larkin and Simon's phrase, PMDs are more *perceptually enhanced* than truth tables. We have also seen that labeling information need not be carried around with each diagram. Such observations suggest that PMDs are a more economical representation that truth tables.

Regarding transferability, more research will be required before any judgment can be made concerning the effective power of PMDs.

It is perhaps the iconic nature of PMDs which most distinguishes them from other representations. In terms of Simon's definitions, PMDs and truth tables are informationally equivalent. However, since they are iconic, PMDs require a simpler form of pattern matching than any purely symbolic representation would require. That is, PMDs make the recognition phase of problem solving much easier.

The iconic nature of PMDs should not be confused with the debate surrounding mental imagery. Much has been written about how problem solving is enhanced by the use of pictorial representations, that is, visual images such as diagrams, photos, maps and internal mental pictures. The diagrams in [LARK87] and [MARK86] are of this type. The computer software Hyperproof (by Barwise and Etchemendy) and Tarski's World both use pictures of blocks to assist in logical proof construction, and various authors have undertaken research into the effectiveness of this graphical approach [STEN92, BARW93, COX93, OBER94]. Much has also been written on the question of how such pictures are actually represented in the brain: whether cognition relies on an inherently imagistic representation or whether the mind uses some propositional (ie language-like or sentential) representation which gives the *illusion* of visual images. See [DENI89] for a detailed analysis of this debate, or [GARD85] for a summary.

Such research lends support to the claim that the visual aspect of PMDs may assist in solving problems in propositional logic. However, PMDs are not visual images in the sense that they look like an actual real-world situation. Rather, than being *realistic* visual images, icons (and hence PMDs) are *symbolic* visual images. They assist memory and problem solving not just in that they are visual, but also in that they act as symbols which represent more information than is explicitly depicted in their visual form.

6.5.5 The Role of Multiple Representations

A major problem when using <u>any</u> representation to teach an abstract concept is that the student may learn the representation but not the underlying concept. As a trivial example, parents may use apples to teach their children that 2+3=5 — "Here are two apples, here are three more, how many have we got when we put them all together?" But whereas the parent imagines the child will learn about numbers, addition and equality, the child may well think that the lesson is about apples.

When given an example of a concept, students will not necessarily abstract from the example that feature which the teacher intends. For instance, when told that affirming the consequent is a logical fallacy, they may falsely infer that all logical fallacies involve conditional statements. However, if several examples of the category or concept are described, the student is more likely to abstract that feature which is common to all the examples: that is, the underlying characteristic which defines the category. Students can be assisted in this process if the examples are chosen to be as diverse as possible so that they have no feature in common apart from the feature which the teacher is attempting to convey.

This principle applies not only to the choices of examples in a categorisation task but also to the choices of representations in any abstract learning task. If only one representation is used, students are quite likely to focus on the representation rather than the underlying concept. They may learn how to use the procedures defined for that representation and may learn how to solve problems within that representational structure, but they may not have learnt the concepts underlying that structure and may be unable to solve conceptually equivalent problems outside that structure. This could be the case regardless of the quality of the representation.

To avoid this, it is helpful to make use of multiple representations. A child should not only be taught that 2+3=5 using apples and oranges and fingers and days, but also using the symbolic form of an arithmetic equation, as well as in unary notation[3] and binary notation. Someone who can manipulate all of these different representations is more likely to have grasped the underlying concepts of numbers, addition and equality. It is even more certain that someone will abstract the underlying concept if they are taught to translate between the various representations: matching the number of apples against fingers, making a notch on a tree as each day passes, converting from decimal to binary, using the ten fingers as binary digits to count up to 1024 etc.

Similarly, when teaching computer science students about circular buffers, they may first learn how to implement them using an array. But if the lesson stops there, they will have learnt something more about arrays, but probably little about circular buffers. They should then learn how to implement a circular buffer using a linked list. In this way they will have seen that a circular buffer is a general concept which is independent of its instantiation in any particular implementation.

[3] That is, what may be called "caveman addition" -- | | + | | | = | | | | |.

In the same way, when teaching logic it is extremely advantageous to teach using more than one representation, and to show how problems and techniques in one representation may be translated into another. (See [PETE94] for some other aspects of multiple representations.)

In this course, various forms of logic are presented, and even when dealing with propositional logic, students are presented with a number of different ways of representing logical expressions and inferences. These include English, set theory, PMDs, truth tables and Lemmon's natural deduction system. The course drills the students in each of these representations, so that they become familiar with the different notations, familiar with the operations allowed within each system, and familiar with the advantages and disadvantages of each representation.

It is hoped that by dealing with this variety of representations, the student will be able to abstract the underlying principles more readily than if they were exposed to any one of the representations. For instance, the concept of a conditional statement is dealt with in each representation as follows:

- English — for instance we note that "Whenever it rains the grass grows" and "If it rains then the grass will grow" are expressions of the same proposition.
- Set theory — the subset relation $A \subseteq B$ means that whenever something is an element of A then it is also an element of B.
- PMDs — material implication is defined by the reverse-and-overlay operation.
- Truth tables — material implication is defined as a particular configuration of T's and F's.
- Natural deduction — the rules MPP (modus ponendo ponens), MTT (modus tollendo tollens) and CP (Conditional Proof) define how material implication can be introduced and eliminated.

Further, the course consistently prompts the students to translate between these representations. Examples and exercises frequently state some propositional expression in one form and require the student to write the same expression in another form (eg to take an argument written in English and to represent it as a propositional sequent). Other questions require the student to prove a propositional sequent in several different ways (eg by a natural deduction proof, by constructing a truth table and by constructing a PMD).

By seeing the same concepts in this variety of representations and by being able to translate between the different representations, students are encouraged not just to learn the individual notational systems and their associated operations, but also to understand the underlying principles of which these representations are but instantiations.

6.5.6 Piagetian Principles of Cognitive Development

It should be noted that much of what is written above differs from a Piagetian analysis. Although the stages of cognitive development suggested by Piaget are a basis for much educational theory for children, they fall short of giving a complete picture, especially for older students.

The Piagetian school plots the development of children through three main stages, the last of which may be further split into two finer divisions [NICK85] —

- Sensori-motor stage (0 to 2 years)
- Pre-operational stage (2 to 7 years)
- Operational stage (7 to 16 years)
    - Concrete Operations stage
    - Formal Operations stage

In the Concrete Operations stage, a student is able to deal effectively with concrete concepts and operations, but not with abstract ones. Because of this, learning is domain-restricted and little transfer can occur. Only in the Formal Operations stage can abstract concepts be learnt and transferred to domains which differ from those in which the concepts were first learnt. For instance, in the Concrete Operations stage, a student may be able to find the union of two finite sets, but not be able to draw a Venn Diagram to represent "some a are b"; a student may be able to construct truth tables involving material implication, but not be able to give a specific example of the fallacy of denying the antecedent[4].

Much of this course in introductory logic requires abstract thinking, even though many concrete examples are used. Since Piaget's stages suggest that the Formal Operations stage should be reached by the age of sixteen, we could assume that university students undertaking this course are capable of learning abstract logical concepts. However, various studies have indicated that a significant percentage of university-aged students have not passed the Concrete Operations stage [NICK85 p 32]. Such students (according to Piaget) will be able to follow and reproduce the examples, but cannot be expected to learn the abstract concepts underlying those examples.

Vygotsky's analysis is at odds with Piaget's on this issue. Piaget suggests that there is no point in presenting information to a person who has not yet reached the stage at which they can assimilate that information. If a student is still in the Concrete Operations stage then teaching must occur through concrete examples and there should be no expectation that abstract learning will occur. In contrast, Vygotsky's model of the zone of proximal development suggests that teaching should always occur in an area *beyond* the student's current capacity. For Vygotksy, it is

---

[4] These examples come from Thornton (1980) quoted in [NICK85 p31].

entirely appropriate to present abstract concepts to students who may only be in the Concrete Operations stage; for how else will they move into the Formal Operations stage?

One of the advantages of using more than one representation is that it promotes abstract learning. If just one representation were taught, then it could be seen by the student as a set of concrete operations to be learnt without any need to look deeper. But if students are exposed to several representational systems they are challenged to grapple with the question of how these systems can represent the same concept when they seem so different. Translating between the representations can be used to prove to the students that the representations are in fact equivalent and leads them to see that the various concrete operations in those different systems share a common yet abstract essence.

Chapter 7: Analysis of Student Responses to Possible Models Diagrams

7.1 Introduction

In the beginning, the research associated with the development of PMDs was of a technical nature, concentrating on the verification of the theoretical correctness of the approach. It was only after the method had been used for two years that consideration was given to the didactical effectiveness of PMDs and the data collected during that early period lacked systematic structure. During the second semester of both 1992 and 1993, the module on logic in the first year computer science course was taught using both truth tables and PMDs (in the manner detailed in the Appendix). During these modules, results from a number of assignments, tests and exams were recorded, but unfortunately some of the original data sources are no longer available. Thus, actual student responses are shown for some of the assessments, while for others there is only summary data.

For comparisons to be made, it would be ideal to teach the same course to two similar groups using PMDs with one group and truth tables with the other. Although this has not been possible, it was possible to teach a short course in which boolean algebra was presented using gates and truth tables. This course was taught during the first semester of 1994 as part of a computer architecture module, to students who would do the full logic course in the following semester. The results of student assessment from this course give some indication of how well students who have never encountered PMDs understand boolean concepts.

In addition, at the end of the 1992 and 1993 logic modules a course evaluation questionnaire was completed to provide students with an opportunity to comment on both the course content and teaching style. These two forms of assessment (assessment *of* students and assessment *by* students) provide the basic data for this chapter.

The data discussed here is typically unstructured and the data collection unsystematic, leading to many uncontrolled variables. Although it is difficult to quantify the data's significance, statistical analysis has been applied to the extent that it is meaningful. Of equal importance is the qualitative aspect of the data. Since the data comes from actual student responses, it has been an invaluable source of assistance in the development and improvement of the teaching methods. The qualitative richness of the data has highlighted various student difficulties and misconceptions, and this has made me better able to address those difficulties.

7.2 Assessment of Students' Competence in Propositional Logic

7.2.1 What Types of Questions should be Asked?

There are three basic concepts related to PMDs (or truth tables) which may be assessed:

- Constructing a PMD (or truth table) for a boolean expression in order to establish whether it is tautologous, contingent or inconsistent;
- Translating between English, truth tables and PMDs; and
- Using PMDs (or truth tables) to establish the validity of a propositional sequent.

7.2.1.1 PMD Orientation

Before any of these concepts can be tested, there is one other question which students should always be asked. It has been noted before that PMDs are generally drawn with unlabeled vertices, and that the ordering of vertices is implicit. Since a common mistake made by students is to order vertices incorrectly, it is essential to require students to draw a labeled PMD before attempting other questions. Very few students have any trouble labeling a PMD with two variables, but a substantial number (approximately 15%) mislabel PMDs with three variables. A complete categorisation of the types of orientation errors is described below, with examples selected from a variety of student responses to exams and assignments.

Consider the following student responses to the task "Draw a fully labeled diagram showing all possible models of three variables, P, Q and R".

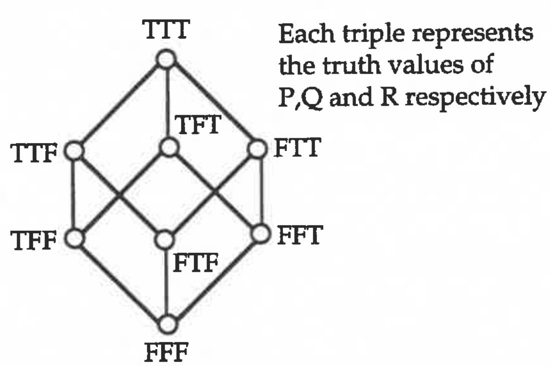


Figure 7.1 — A Correctly Labeled PMD for Three Variables

Figure 7.1 shows the correct answer, complete with a brief note explaining the labeling convention. Notice that if this graph is visualised as a cube (which students are encouraged to do), then the cube has a face on which each corner has P with the value "T" and an opposing face on which each corner has P with the value "F" and that there are similar faces for Q, ~Q, R and ~R.

The following Figures show a variety of incorrect student responses to the same task.

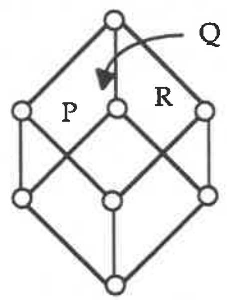


Figure 7.2 — Incorrect — Vertices Not Labeled

In Figure 7.2, the student has indicated the three faces for P, Q and R, but has not labeled the vertices. This raises one of the unfortunate short-comings of written exams as a form of assessment. It may well be that the student mis-understood what was required by the question and if asked to expand or clarify her answer she may have been able to give the correct answer.

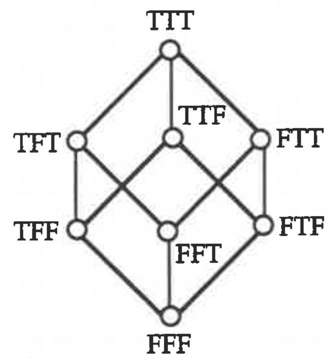


Figure 7.3 — Incorrect — Non-standard Orientation

Figure 7.3 shows another student's response to the same question. In this case the labels indicate a distinct P-face, Q-face and R-face, but the orientation of the imagined cube differs from the convention. As long as it is used consistently, this is a reasonable alternative. However, the convention is important in that it allows anyone familiar with PMDs to read unlabeled diagrams without ambiguity and since the convention has been violated, this answer would only gain half marks.

Requiring the student to draw a fully labeled diagram alerts the marker to the student's non-standard orientation. The marker can then interpret answers to other questions with this orientation in mind. If students were not required to initially draw a labeled PMD, then with later questions it would be impossible for the marker to differentiate between logically flawed answers and answers which are incorrect but understandable when viewed from an alternate orientation.

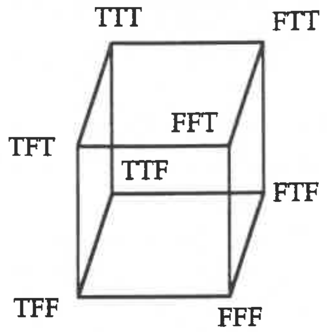


Figure 7.4 — Incorrect — Non-standard Orientation

Figure 7.4 shows another example of non-standard orientation. It is clear from the way the diagram is drawn that the student is visualising a cube. Only one student (out of roughly 100 students who have been exposed to this course) drew the graph in this way.

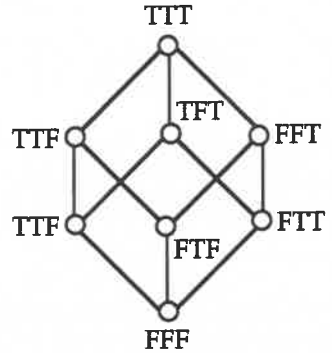


Figure 7.5 — Incorrect — Various Incorrect Labels

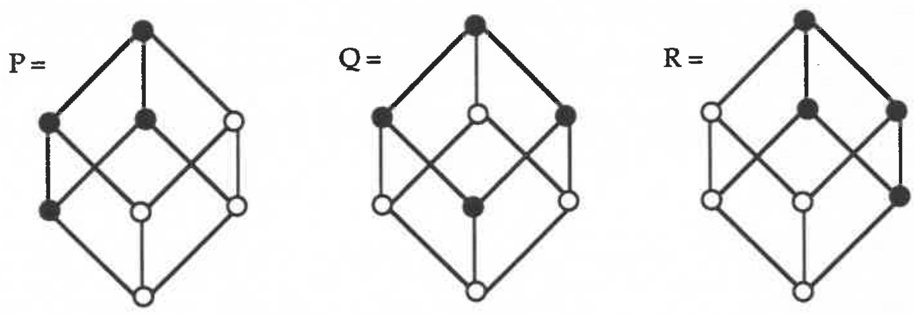


Figure 7.6 — Correctly Labeled Faces

Figure 7.5 has one label repeated (TTF) and two labels have exchanged position (FFT and FTT). I believe this student had grasped the concept of separate faces for P, Q and R (since in an earlier assessment he had correctly drawn Figure 7.6) but became confused in the translation to a labeled graph.

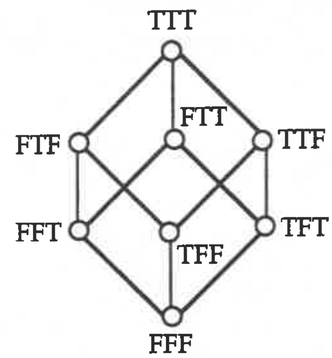


Figure 7.7 — Incorrect — No Apparent Order

The comment was made earlier that Figure 7.3 is not so much a mistake of *logic* as a mistake of *communication*. But Figure 7.7 is fundamentally flawed as it has neither a P, Q nor R face. The student has listed all the correct labels, but all except two are in the wrong position.

The visual rules for manipulating PMDs are such that even the confused labeling of Figure 7.7 would work, if it were used consistently. This remains true for any labeling pattern as long as all eight triples TTT, TTF, TFT, FTT, TFF, FTF, FFT and FFF are used. However, unless some clear order is maintained the confusion will make the possibility of consistency very unlikely.

7.2.2 Class Test 1992

The Mid-semester Test in 1992 included the following questions:

1. Draw a fully labeled diagram showing all possible models of two variables, P and Q.
2. For each of the models shown in 1, write English interpretations of P and Q which fit the model.
3. Draw a separate fully labeled diagram showing all possible models of three variables, P, Q and R.
4. Draw either a possible models diagram or a truth table for each of the following wffs. Classify each wff as either contingent, tautologous or inconsistent.
   i) (P⇒Q)⇒(Qv~P)
   ii) ~(PvQ)&(PvQ)
   iii) (P&~R)⇒Q

The results of fifty students' answers to Question 4 allow a comparison of PMDs with truth tables. The percentages of correct and incorrect results are shown in Table 7.1[1]. Note that in this question students could choose whichever method they preferred.

[1] In all the Tables in this chapter I have specified the number of students who sat for the exam/test or submitted an assignment, and then shown the percentage of students falling into

Table 7.1 — Class Test 1992 Question 4

	n=50	(P⇒Q)⇒(Qv~P)	~(PvQ)&(PvQ)	(P&~R)⇒Q
PMD	Correct	80	78	30
	Incorrect	10	10	22
Truth Table	Correct	6	10	12
	Incorrect	4	2	32

From these results it can be seen that when dealing with boolean expression containing two variables, 89% of the students chose to use PMDs rather than truth tables. The odds of answering the question correctly using PMDs was 7.90, compared to 2.67 for truth tables[2].

When faced with an expression containing three propositional variables, a larger proportion opted to use truth tables, but the majority (52%) still used PMDs. Once again, the odds of giving a correct answer using PMDs (1.36) was much higher than that for truth tables (0.38). (Note that 4% did not even attempt this three-variable question.)

The overall sample size here (and in the cases which follow) is relatively small, however, a more important restriction on any statistical analysis is the low numbers of responses in some categories. For instance, only one student submitted an incorrect truth table for ~(PvQ)&(PvQ). Nevertheless, we can calculate the odds of getting a correct answer with one method as compared to the other, and we can apply a χ^2 test to determine whether the difference is significant.

If the three questions are analysed separately, the first two show an Odds Ratio in favour of PMDs, but the cohort sizes are too small for this to be judged significant. In the third case the choice between PMDs and truth tables was more evenly balanced. In this case, the Odds Ratio is 3.64 indicating that a student choosing to use PMDs was 3.64 times more likely to give a correct answer than a student choosing to use truth tables. The Mantel-Haenszel Summary Chi Square value is 4.39, with a P value of 0.036 (which is within the standard 0.05 significance level). In summary, although there is no significant difference indicated between the choice of PMDs and

---

appropriate categories. However, the percentages in the table occasionally add up to less than 100 since not all students attempted all questions. It cannot be assumed that failure to attempt a question means the students was incapable of answering that question: it may be that time restrictions made them choose to concentrate on other areas of the assessment instead. There are also tables where rounding errors lead to totals of slightly over 100.

[2] The "odds" of some event is the number of times that event occurs (or is expected to occur) divided by the number of times it does not occur (or is not expected to occur). Thus if 15 students correctly construct a PMD for (P&~R)⇒Q, and 11 students attempt to but fail, then the odds of succeeding is 15:11 or 1.36. I originally quantified such comparisons as error rates: an error rate of 11 out of 26, or 42% in the previous example. However, the odds measurement is a more useful figure for some of the other statistical tests. The term "Odds Ratio" refers to a comparison of the odds for two different methods.

truth tables for wffs with two propositional variables, there is a significant difference for the wff which had three variables.

In order to increase the sample size, the three questions may be combined. When this is done, the Mantel-Haenszel Weighted Odds Ratio is 3.39 in favour of PMDs. The Mantel-Haenszel Summary Chi Square value is 5.38, with a P value of 0.020 (which is well within the standard 0.05 significance level).

This analysis, based on the Odds Ratio and the χ^2 test, relies on the assumption that the choice of method (ie PMD or truth table) is random. This assumption would be justified if we could be assured that the quality of students choosing one method was the same as the quality of students choosing the other. But what if the better students chose one method and poorer students the other? In similar situations in later sections a Logistic Model can be build to allow for this, but in the current instance the raw data necessary for such analysis is no longer available.

7.2.3 Final Exam 1992

The Final Exam in 1992 included the following question:

> For each of the following wffs construct both a possible models diagram and a truth table. State whether the wffs are tautologous, contingent or inconsistent. (Please indicate whether you constructed the possible models diagram first or the truth table first.)
>
> 1. P⇒((~Qv~P)⇒P)
> 2. ((PvQ)&R)⇒~ Q

In this question, *both* a PMD and a truth table must be constructed. However, the task of translating between one and the other is reasonably trivial (see Section 5.4.2) and so it may be assumed that a student would first construct the one which they consider easiest and then construct the other by translating the first. The results are summarised in Table 7.2.

Table 7.2 — Final Exam 1992

		n=55	P⇒((~Qv~P)⇒P)	((PvQ)&R)⇒~ Q
PMD first	PMD	Correct	62	45
		Incorrect	24	18
	Truth Table	Correct	51	35
		Incorrect	27	22
TT first	Truth Table	Correct	11	20
		Incorrect	4	11
	PMD	Correct	15	15
		Incorrect	0	15

With both two- and three-variable expressions the majority of students preferred to construct the PMD before the truth table (75% for two-variables and 67% for three-variables). The odds of success for those opting to construct PMD first was 2.55 compared with 1.76 for those opting for truth table first. Once again a small proportion (5%) did not attempt the harder three-variable question.

The structural complexity of this data make it difficult to analyse in any more detail and no meaningful conclusions can be drawn. If we were to simply focus on the method they chose to use first and compare the odds of success, the difference is insignificant.

7.2.4 Class Test 1993

The Mid-semester Test in 1993 included the following questions:

1. Draw a fully labeled diagram showing all possible models of three variables P,Q and R.
2. Construct both a PMD and a Truth Table for the wff ~(P v (P ⇒ (Q&P))). Is this wff tautologous, contingent or inconsistent?
3. Construct either a PMD or a Truth Table for the wff (Qv~R) & (Q⇒P). Is this wff tautologous, contingent or inconsistent?

The results for Question 2 (where both PMD and truth table were required) are summarised in Table 7.3. As with previous results, the odds of success for PMDs (5.00) is higher than for truth tables (1.56).

Table 7.3 — Class Test 1993 Question 2

n=42		~(P v (P ⇒ (Q&P)))
PMD	Correct	83
	Incorrect	17
Truth Table	Correct	60
	Incorrect	38

The results for Question 3 (in which students could choose which to construct) are summarised in Table 7.4.

Table 7.4 — Class Test 1993 Question 3

n=41		(Qv~R) & (Q⇒P)
PMD	Correct	27
	Incorrect	17
Truth Table	Correct	24
	Incorrect	32

These results indicate that for an expression with three-variables, more students (56%) prefer to use a truth table, but those who do so are less likely to do so correctly (odds of 0.77) compared with those who use a PMD (1.57).

If the two questions are analysed separately, Question 2 shows an Odds Ratio of 3.2 in favour of PMDs, and the Mantel-Haenszel Chi Square value is 5.11, with a P value of 0.023 (which is within the standard 0.05 significance level). In the case of Question 3 the Odds Ratio is 3.64 in favour of PMDs, but the Mantel-Haenszel Chi Square value of 1.23 is not significant.

When the two questions are combined, the Mantel-Haenszel Weighted Odds Ratio is 2.68 in favour of PMDs and the Mantel-Haenszel Summary Chi Square value is 5.16, with a P value of 0.023 (which is well within the standard 0.05 significance level).

7.2.5 Assignment 1993

During the 1993 module, students were required to hand in a written assignment which included the following questions:

1. At the bottom of page 48 of the Course Notes [see Appendix, page A–48], four different ways of proving a propositional sequent are described (note that method ii is actually two methods — PMD and Truth Table). Prove both of the following sequents by <u>each</u> of the four methods. In both cases, state which of the four methods seemed easiest to you.
   i) $\sim P \vee Q, \sim Q \Rightarrow P \;\therefore Q$
   ii) $P \Rightarrow (Q \Rightarrow R), \sim R \;\therefore P \Rightarrow \sim Q$
2. Show that the following argument is invalid —

   If Alice wins first prize, then Bob wins second prize and if Bob wins second prize then Carol is disappointed. Either Alice wins first prize or Carol is disappointed. Therefore Bob does not win second prize.

The results of the analysis of answers to Question 1, categorised into the four methods are summarised in Table 7.5.

In the two-variable case, the odds of getting the answer correct with each of the four methods are 0.48, 2.45, 2.45 and 17.50 respectively. In the three-variable case, the odds are 4.43, 1.11, 1.31 and 3.38 respectively. It is difficult to infer anything from these figures regarding the deductive proof method in general, since the difficulty of constructing such proofs is very dependent on the semantic content of the sequent. In contrast, the construction of truth tables and PMDs is a mechanical procedure for which the complexity depends solely on the number of distinct variables and the number of operators.

Table 7.5 — Assignment 1993 Question 1

	n=38	$\sim P \vee Q, \sim Q \Rightarrow P \;\therefore Q$	$P \Rightarrow (Q \Rightarrow R), \sim R \;\therefore P \Rightarrow \sim Q$
Deductive proof	Correct	32	82
	Incorrect	66	18
Tautology by TT	Correct	71	53
	Incorrect	29	47
Tautology by PMD	Correct	71	55
	Incorrect	29	42
PMD subset	Correct	92	71
	Incorrect	5	21

McNemar's Test was used to analyse this data further. McNemar's Test applies to one pair of methods, and so in this situation it must be repeated for each possible pairing of the four methods. For each of these pairings, the number of students getting the answer correct with one method but incorrect with the other method was counted. Let r_1 be the number of students who answered the first correctly, but the second incorrectly, and let r_2 be the number of students who answered the second correctly, but the first incorrectly. If the two methods were equally difficult, we would expect r_1 to be the same as r_2. But if —

$$\left[\frac{\dfrac{r_1}{r_1+r_2}-\dfrac{1}{2}}{\sqrt{\dfrac{1}{4.n}}}\right]^2 > 3.84$$ (the χ^2 threshold for the 0.05 significance level),

then the method corresponding to the greater of r_1 and r_2 must be significantly better than the other.

In the two-variable case, McNemar's Test shows that the deductive proof method is significantly worse than all the others, and that the PMD subset method is significantly better than all the others. In the three-variable case, the deductive proof method is shown to be significantly better than both truth table and PMD tautology methods, and the PMD subset method is significantly better than the truth table method.

Fifteen of the 38 students commented on which method they found easiest and the frequencies of these preferences are recorded in Table 7.6. One can see from this table the pattern noted previously: namely that while PMDs are preferred for expressions with two variables, they are not so popular for expressions with three variables. One point of interest is that two students who declared that truth tables were easiest nevertheless constructed incorrect truth tables.

If more than three variables were required it would be expected that virtually all students (and practiced logicians as well) would prefer to build some type of deductive proof rather than any of the other methods.

Table 7.6 — Assignment 1993 Question 1 Preferences

n=15	~P v Q, ~Q⇒P ∴ Q	P⇒(Q⇒R), ~R ∴ P⇒~Q
Deductive proof	7	40
Truth table	33	47
PMD	27	7
PMD subset	33	7

Question 2 required the student to first translate the argument from English into a propositional sequent, and then to use one of the methods from Question 1 to disprove it. Of the 35 students who attempted this question, 20% did not correctly express the task in symbolic notation — ie they did not solve the translation phase. Table 7.7 shows which of the four methods of proving validity were used by students in order to answer Question 2.[3]

Table 7.7 — Assignment 1993 Question 2

n=32		Alice, Bob and Carol
Deductive proof	Correct	N.A.[4]
	Incorrect	16
Tautology by TT	Correct	44
	Incorrect	9
Tautology by PMD	Correct	16
	Incorrect	0
PMD subset	Correct	16
	Incorrect	0

The majority of students chose to test the sequent by truth table, which is what we now expect since the expression has three variables. But note again that those who use either of the methods involving PMDs have a lower error rate (in this case 0%). The fact that most students answered this question correctly places a severe limit on the significance of any statistical analysis.

A Logistic Model was constructed to correlate the quality of the students with the probability of solving the problem correctly using the various methods. The total mark for the assignment (towards which this question contributed 10%) was used as an indicator of student quality. This mark by itself was not significantly correlated to the ability to solve this question correctly (the model gave a deviance of 2.10 with one degree of freedom, whereas a result of greater than 3.84 is required for significance at the 0.05 level). Similarly, the method chosen to solve the problem was not a significant predictor of success[5] (the model gave a deviance of 3.78 with two degrees of freedom, whereas a result of greater than 5.99 is required for significance at the 0.05 level).

The Logistic Model provides one other useful piece of information. In an earlier section I raised the question of whether the quality of student might effect the choice of method. The model here indicates that when interaction between student quality and choice of method is allowed,

[3] The value of n=32 for the results shown in Table 7.7 does not exactly match the 28 students who correctly translated the English into symbolic notation. The reason for this is that some students who correctly translated the argument did not attempt to invalidate it, and some students who mis-translated the argument nevertheless proceeded to correctly invalidate the sequent they had (mis-)constructed.

[4] Since it is impossible to construct a proof by natural deduction of the <u>invalidity</u> of a sequent, any attempt to do so was automatically marked as incorrect.

[5] A perfect prediction could be made in this question if the method chosen was the construction of a deductive proof. But this is only the case trivially since any attempt at such a proof is guaranteed to fail, and so that method was not included in the Logistic Model.

the correlation with success has a deviance of 0.0001. That is, this data gives no evidence at all that the effect of choice of method depends on student quality.

7.2.6 Final Exam 1993

The Final Exam in 1993 included the following questions:

1. Draw a fully labeled Possible Models Diagram for two variables and a separate fully labeled Possible Models Diagram for three variables
2. Describe the four ways by which a propositional sequent may be proved valid.
3. Use each of the above four methods to prove the sequent $Pv(\sim P \Rightarrow Q), \sim Q \therefore P$
4. Convert the following argument into propositional logic notation. Show whether the argument is valid or invalid.

   > It is a well established principle that if you have well-designed software and well-designed hardware, then you will get an efficient computing system. One can also assume that if software is not well-designed then the computing system will not be efficient. Therefore, whenever you have well-designed hardware it implies that the software is also well-designed.

5. Show whether the following wffs are equivalent using either Possible Models Diagrams or Truth Tables.
   - $\sim(PvQ) \equiv \sim P \ \& \sim Q$
   - $P \Rightarrow (QvR) \equiv (P \Rightarrow Q)vR$

The results of the analysis of answers to Question 3 are summarised in Table 7.8. The high error rate seems to indicate that students found this to be a difficult question.

Table 7.8 — Exam 1993 Question 3

n=45		Pv(~P⇒Q), ~Q ∴ P
Deductive proof	Correct	56
	Incorrect	36
Tautology by TT	Correct	53
	Incorrect	38
Tautology by PMD	Correct	58
	Incorrect	27
PMD subset	Correct	56
	Incorrect	18

The odds of success using the four methods are 1.56, 1.41, 2.17 and 3.13 respectively, showing (among other things) that using the PMD subset method is 2.22 times more likely to be applied correctly than the truth table method.

When McNemar's Test was applied to the data from Question 3, the PMD methods consistently performed better than the other methods, but not to the required level of significance. One problem with applying McNemar's Test to this data is that sixteen of the 45 students did not attempt to use all four methods. Since we can not assume that failure to use a method implies inability to use that method, a lot of potential pairwise comparisons were lost.

Question 4 required the student to first translate the argument from English into a propositional sequent, and then to use one of the methods from Question 2 to disprove it. Of 43 students who attempted this question, all but two correctly handled the translation phase. Five students who did perform the translation correctly did not attempt any form of disproof. Table 7.9 shows which of the four methods of proving validity were used by the remaining 38 students.

Table 7.9 — Exam 1993 Question 4

n=38		Hardware/Software
Deductive proof	Correct	N.A.[6]
	Incorrect	16
Tautology by TT	Correct	21
	Incorrect	24
Tautology by PMD	Correct	13
	Incorrect	5
PMD subset	Correct	16
	Incorrect	5

In this examination the students showed a distinct preference for PMDs, even when the propositional expression contained three variables (both here and in Question 5). The odds of getting the answer correct using the four methods was zero for deductive proof, 0.89 for tautology by truth table, 2.50 for tautology by PMD and 3.00 for the PMD subset technique. The differences between these odds are not significant.

If one combines the two PMD methods, then the Odds Ratio of 3.09 shows that students choosing to use PMDs are much more likely to answer correctly than students choosing to use truth tables. Applying a χ^2 test to establish the significance of this difference is not valid in this case due to the small number of students who used PMDs incorrectly.

A Logistic Model for the results of Question 4 (similar to that described for Assignment 1993 Question 2) revealed that neither student quality, choice of method, nor the interaction of the two were significant predictors of success.

Question 5 required the students to analyse four different wffs. Since the question allowed students to choose whichever method they preferred, the results shown in Table 7.10 may be used to compare their preferences.

---

6 Since it is impossible to construct a proof by natural deduction of the invalidity of a sequent, any attempt to do so was automatically marked as incorrect.

Table 7.10 — Exam 1993 Question 5

n=46		~(PvQ)	~P & ~Q	P⇒(QvR)	(P⇒Q)vR
PMD	Correct	76	80	57	57
	Incorrect	13	9	9	9
Truth Table	Correct	11	11	30	30
	Incorrect	0	0	2	2

Although students had to construct truth tables or PMDs for all four of these wffs, the question clearly consisted of two parts: a comparison of the wffs ~(PvQ) and ~P&~Q; and a comparison of the wffs P⇒(QvR) and (P⇒Q)vR. In answering the first part, there were two students who constructed an incorrect PMD for ~(PvQ) but a correct one for ~P&~Q. In all other cases, students either analysed both wffs correctly or both incorrectly. One notable feature of this question was that three students gave incorrect answers to the two-variable expression yet gave the correct answer for the supposedly harder three-variable expression.

In both two- and three-variable cases, the majority of students chose to use PMDs rather than truth tables, though the difference was less in the three-variable case. The error rate for those who chose PMDs was 13%, notably higher than the error rate for truth tables (5%). An analysis of the Odds Ratios is again difficult because some of the cohort sizes are so small (including zero in some cases). When all four columns in Table 7.10 are analysed together, the Mantel-Haenszel Weighted Odds Ratio is 0.32 (ie those who used PMDs had 0.32 times the probability of succeeding than those who chose to use truth tables). However, the Mantel-Haenszel Summary Chi Square value of 1.27 has a P value of 0.26, indicating that this measure has little significance.

In order to allow for differences in student quality, we again constructed Logistic Models: one for the two-variable cases and another for the three-variable cases. The two-variable model showed that student quality *is* a good predictor of success (the deviance of 9.38 with one degree of freedom well exceeds the 0.05 significance threshold of 3.84), but choice of method is not a good predictor of success (the probability of success favoured the PMD method, but with a deviance of only 1.52), nor is the interaction between student quality and choice of method significant (deviance 0.0002).

The three-variable model showed that neither student quality nor choice of method were significant predictors of success (deviances of 1.33 and 1.75 respectively). However, the interaction between student quality and choice of method was significant (deviance of 6.44). In this case, the odds favoured the PMD method and showed that as student quality increased, the probability of succeeding using PMDs increased at a faster rate than the probability of succeeding using truth tables. In other words, students of poorer quality performed equally as poorly (or nearly so) with both methods, whereas better students were more likely to answer correctly if they used PMDs.

7.2.7 Truth Tables Only (1994)

During the first semester of 1994, a four week course in boolean algebra and logic circuits was taught. In this module PMDs were never mentioned but truth tables were used extensively. Since the emphasis in this course was algebraic rather than propositional, the notation here differs from that used in the rest of this thesis. The symbols 0 and 1 are used instead of T and F; disjunction is indicated by addition ("+"); conjunction by multiplication ("." or simply concatenation); and negation by either an overbar (eg $\overline{A+B}$) or a prime (eg (x+y)′). There is no reason to believe that this notation is any harder or easier to master than the notation used by students in the courses discussed previously.

A class test at the end of this section included the following questions —

1. Draw truth tables to show whether or not the following expressions are equivalent —
   i) $\overline{A+B} \equiv \overline{A}.\overline{B}$
   ii) (x+y)′(x+y) ≡ xyx′
2. Draw a truth table for the expression x(y+z′)+x′yz.

These questions require the student to construct a total of five truth tables: four for wffs with two variables and one for a wff with three variables. In the context of boolean algebra and logic circuits, it is not appropriate to introduce material implication, and so these questions only involve the operators negation, conjunction and disjunction. The results of the analysis of answers to these questions are summarised in Table 7.11.

Table 7.11 — Truth Tables Only (1994)

n=54	$\overline{A+B}$	$\overline{A}.\overline{B}$	(x+y)′(x+y)	xyx′	x(y+z′)+x′yz
Correct	85	89	72	74	56
Incorrect	13	9	20	22	37

The low success rate for the wff containing three variables indicates that students find truth tables for such wffs more difficult than truth tables for wffs with only two variables. This is consistent with what we have noted before about such wffs, and applies to both truth tables and PMDs.

As with Question 5 in the 1993 Final Exam, Question 1 required pairs of wffs to be compared. In answering Question 5 in the 1993 Final Exam, it was noted that very few students would analyse one wff correctly but the other incorrectly. However, in Question 1 in 1994, 13 students (24%) made this type of mistake.

The wffs $\overline{A+B}$ and $\overline{A}.\overline{B}$ correspond to the wffs ~(PvQ) and ~P&~Q which featured in Question 5 of the Final Exam of 1993. The wff (x+y)'(x+y) corresponds to ~(PvQ)&(PvQ) which featured in Question 4 of the Class Test of 1992. Because of this, a comparison of students' answers to these questions in 1992, 1993 and 1994 may suggest whether teaching both PMDs and truth tables yields any improvement over teaching just truth tables. In the first place we will ignore the PMD results and simply ask "Does a course which includes both truth tables and PMDs make students more capable of *constructing truth tables* when compared to students in a course which only covers truth tables?"

The success rates for construction of truth tables for $\overline{A+B}$ and $\overline{A}.\overline{B}$ in 1994 (from Table 7.11) could be compared with the success rates for construction of truth tables for ~(PvQ) and ~P&~Q in 1993 (from Table 7.10). However, the entries of zero in Table 7.10 mean that no Odds Ratio can be calculated, and the Mantel-Haenszel Summary Chi Square figure shows that there is no significant difference between the data for the two years.

A similar comparison could be made between the success rates for construction of a truth table for (x+y)'(x+y) in 1994 (from Table 7.11) and for ~(PvQ)&(PvQ) in 1992 (from Table 7.1). The Odds Ratio is 0.71 (in favour of the 1992 results) but the Fisher Exact test shows this to be insignificant (two-tailed P-value of 1.0).

When the figures from the previous two paragraphs are combined, the Mantel-Haenszel Weighted Odds Ratio is 0.35 in favour of the courses in which PMDs were taught, however, this is also indicated to be insignificant by the low Mantel-Haenszel Summary Chi Square figure (0.38).

Thus, the data leans towards an affirmative answer to the question "Does a course which includes both truth tables and PMDs make students more capable of *constructing truth tables* when compared to students in a course which only covers truth tables?", but not to a sufficiently significant extent. We could look at the data from a slightly different angle and ask "Does a course which includes both truth tables PMDs make students more capable of *correctly analysing wffs* when compared to students in a course which only covers truth tables?" This perspective treats the 1994 results as a baseline (when truth tables were the only tool available), and allows the comparison with previous years to include the PMD data.

However, even taking this broader perspective yields unconvincing results. In each case where 1994 results are compared with 1992 or 1993 results, the Odds Ratio favours the courses which taught PMDs, but not to a significant degree. When all the data is combined the Mantel-Haenszel Weighted Odds Ratio is 0.72 (in favour of the pre-1994 courses), but the Mantel-Haenszel Summary Chi Square value of 0.58 (with a P value of 0.44) shows that the difference between the 1994 data and the earlier data is not significant.

7.3 Discussion of Quantitative Analysis

7.3.1 Summary of Statistically Significant Results

The variety in the structure of the available data has required that a number of different statistical methods must be used. Despite this, a clear general picture emerges in which the teaching of PMDs is shown to be beneficial with respect to the following:

- In the Class Test of 1992, the answers to Question 4 were more likely to be correct when the student used PMDs rather than truth tables. Although the advantage was insignificant for the simple two-variable cases, it was significant when the wff contained three distinct variables.

- In the Class Test of 1993, students were able to analyse two- and three-variable wff more reliably with PMDs than truth tables. The overall advantage was statistically significant, although not for the three-variable case.

- In the Assignment of 1993, students who used the PMD subset method to check the validity of a sequent in Question 1 performed better than students using the other methods. The difference was significant for both two- and three-variable cases. Student's ability in Question 2 showed no significant difference between the methods.

- In the Final Exam of 1993, the results of Questions 3 and 4 favour PMDs, whereas the results of Question 5 favour truth tables. However, in all three questions the difference was not significant.

- In 1994 the students learnt truth tables but not PMDs. If their results are compared with those of previous students who have learnt both methods, we find that exposure to PMDs increases the ability to correctly analyse two-variable wffs. Here again the difference is not significant.

In summary, there is reasonable evidence to suggest that students perform better when using PMDs than when using truth tables. In some cases there is statistically significant data in favour of this assertion, and in no case has there been statistically significant data to the contrary. Furthermore, there is a strong tendency among students to use PMDs rather than truth tables for situations involving two distinct propositional variables when given the choice, although this tendency is reduced when the number of propositional variables increases to three.

Further study may clarify why individual students prefer one method over the other. It may be conjectured that, when given a choice, students will use the method they find easiest. At least some of the data (the low error rates in Table 7.10 for instance) may be explained by hypothesising that students have learnt which method suits them and hence have chosen the method they can work with correctly. Perhaps if students were made aware of the fact that the odds of giving a correct answer is higher when using PMDs, an even greater proportion may choose to use them rather than truth tables. It is perhaps also the case that some students avoid using PMDs for the three-variable problems because of a difficulty with three-dimensional visualisation.

Even though there are indications that students prefer PMDs over truth tables and that those using PMDs are more likely to give correct answers, there are reasons to treat these results with caution.

7.3.2 Reasons to Doubt the Reliability of these Results

Various comments have been made already about the unsystematic way in which the data was gathered. The history of this research gives rise to a number of issues which may have biased the results and their subsequent analysis.

- Many of the cohort sizes were too small to allow any convincing conclusions to be drawn.
- In several cases, the original data was no longer available. This forced the statistical analysis to be based on summary data and limited the range of statistical tools which could be used.
- The data was based on questions which were designed for assessment of student ability rather than assessment of PMDs or of teaching methods.
- The research and analysis has been carried out by the same person that invented the PMD approach and who taught all the courses. Significant bias towards PMDs may have been introduced by the lecturer's enthusiasm for the PMD approach.
- In many cases it is not clear which variable is being measured by the statistical tests. Although the intended variable under study is the students' ability using various logical methods, students' overall ability and the students' preferences may have interfered. In some cases a Logistical Model was used to allow for such variables, but this was not always possible.
- Whereas the logic courses in 1992 and 1993 were full thirteen week modules, the 1994 course (in which only truth tables were taught) lasted only four weeks.

7.4 Some Qualitative Analysis

When working with truth tables, at least three students drew the PMD for material implication in the corner of the page. It can safely be assumed that this doodle is a memory-aid to help them deal with material implication correctly in the truth table. One student also doodled the PMDs for disjunction and conjunction in the margin next to answers to truth table questions. Incidents such as these add weight to the assertion that teaching PMDs is of more benefit to students with respect to their ability to solve problems in propositional logic than just teaching truth tables.

There is a variety of such qualitative data which deserves discussion, most of which involve observations of mistakes made by students. By noting such mistakes, one can attempt to infer what misconception in the student's mind led to that mistake. A categorisation of the common logical and procedural mistakes indicates which aspects of teaching should be stressed or improved. From the various assessments of students described above, the following categorisation can be proposed.

First, there are several recurring errors which led to incorrect truth tables and/or PMDs, but which were neither errors computing truth table values nor errors manipulating PMDs. In each of the tests and assignments whose results are described above, between 25% and 50% of the mistakes related to framing the answer to the question incorrectly. For instance, when asked to prove the sequent Pv(~P⇒Q), ~Q ∴ P, a student could convert this sequent to the corresponding conditional (Pv(~P⇒Q))⇒(~Q ⇒ P) and show that expression to be tautologous using either truth table or PMD. Many students made mistakes in the conversion phase, before either truth table or PMD was used. This perhaps indicates a need to place a greater emphasis on this conversion process in future courses.

As with any mental procedure there is also a trivial-slip-up factor included in the error rates. Up to 10% of errors with both truth tables and PMDs could be accounted for by one-off accidents caused by exam pressures rather than misconceptions. Such accidents are indicated when a student makes a mistake but fails to repeat the mistake in similar situations. Another 10% of the errors were too confusing to interpret.

The errors remaining after discounting the above, are described below in roughly decreasing order of frequency.

7.4.1 Typical Errors with Truth Tables

1. T⇒F: When the antecedent is true and the consequent false, the overall value of a material implication is defined to be false. But students may mistakenly say that such a case yields true. This is perhaps the most disturbing mistake since this case signifies the very essence of material implication.

   Consider Figure 7.8 for example, which shows a student's answer to Question 3 in the 1993 Final Exam. Values in this truth table are all correct except for those indicated by an asterisk (*). On the fourth row there are two instances where the student has mistakenly indicated that T⇒F yields the value T. (The first asterisk on that row of the truth table follows on from these two other mistakes and is not in itself a mistake.) The student has used different colours to indicate the order in which the operators were processed, and in this she is quite correct. Her only mistake is in the processing of material implication.

P	Q	(P	v	(~P	⇒	Q))	⇒	(~Q	⇒	P)
T	T	T	T	F	T	T	T	F	T	T
T	F	T	T	F	F*	F	T	T	T	T
F	T	F	T	T	T	T	T	F	F*	F
F	F	F	T*	T	T*	F	T	T	T*	F

Figure 7.8 — Incorrect — T⇒F Yields the Value T

Another student constructed the truth table in Figure 7.9 while answering Question 3 in the 1993 Final Exam. Once again asterisks indicate values in the table which are incorrect. In the fourth row, the first error suggests that the student believes that T⇒F yields T. (The second mistake in the fourth row is simply a consequence of the first.) In this case, the student did not construct any other truth tables in this exam and we cannot discount the possibility that this mistake was accidental.

P	Q	~P	~P⇒Q	Pv(~P⇒Q)
T	T	F	F*	T
T	F	F	T	T
F	T	T	T	T
F	F	T	T*	T*

Figure 7.9 — Incorrect — T⇒F Yields the Value T

Yet another student seems to have made this same error in his answer to Question 3 in the 1993 Final Exam. The truth table shown in Figure 7.10 shows three values in the fourth row to be incorrect. These could be accounted for by the hypothesis that the student thought that T⇒F yields and overall value of T. (In this case it may also be that the student filled the whole table with T's for a laugh.)

P	Q	(P	v	(~P	⇒	Q))	⇒	(~Q	⇒	P)
T	T		T		T		T		T	
T	F		T		T		T		T	
F	T		T		T		T		T	
F	F		T*		T*		T		T*	

Figure 7.10 — Incorrect — T⇒F Yields the Value T

2. F⇒T: When the antecedent is false and the consequent true, the overall value of a material implication is defined to be true. But students may incorrectly judge this to be false.

The example previously shown in Figure 7.9 shows this mistake in the first row. Two other students constructed exactly the same truth table for ~P⇒Q. It may be that these students have internalised an incorrect truth table as a definition of material implication in which the rows T⇒F and F⇒T have been reversed (Figure 7.11b rather than the correct 7.11a). At least two other students who sat for the 1993 Final Exam exhibited errors which could be accounted for by this hypothesis.

P	Q	P⇒Q
T	T	T
T	F	F
F	T	T
F	F	T

7.11a

P	Q	P⇒Q
T	T	T
T	F	T
F	T	F
F	F	T

7.11b

Figure 7.11 — Correct (a) and Incorrect (b) Truth Tables Defining Material Implication

3. F⇒F: When both antecedent and consequent are false, the overall value of a material implication is defined to be true. But students may incorrectly judge this to be false.

The example given previously in Figure 7.8 also shows this mistake. The errors in the second and third rows of the truth table can be explained by hypothesising that the student believes that F⇒F yields an overall value of F. In answer to Question 4 in the same exam this student consistently showed this same mistake.

4. Confusing columns: When filling in a truth table, it is important to fill in columns in the correct order so that the boolean operators are processed in the correct precedence sequence. But some students become confused about which columns are being applied to which operator. For instance, consider the incorrect truth table in Figure 7.12.

P	Q	P	⇒	(~	P	v	Q)
		1	2	3	4	5	6
T	T	T	F	F	T	T	T
T	F	T	F	F	T	F	F
F	T	F	T	T	F	T	T
F	F	F	T	T	F	T	F

Figure 7.12 — Incorrect Truth Table for P⇒(~PvQ)

The numbers 1,2,3,4,5 and 6 will be used to refer to the columns of this truth table in order to discuss the order in which they should be filled them in. Columns 1,4 and 6 are simply repeats of the columns representing P and Q on the left hand side. These columns should be filled in first (in any order) or they may be left out completely by confident students. After that, it is crucial that the remaining columns be filled out in correct order: column 3 based on the values in column 4; then column 5 based on columns 3 and 6; and finally column 2 based on columns 1 and 5.

In Figure 7.12, however, the order of calculation is incorrect. Everything is correct up until the values for column 2 are to be calculated. But instead of determining the truth values for column 2 on columns 1 and 5, the calculation is mistakenly based on columns 1 and 3. In this case three of the entries in column 2 remain correct, but a "F" is entered in the first row instead of the correct "T".

The example in Figure 7.12 is an invented one. It would be unusual for a student to make this type of mistake in such a simple wff, but for a longer expression, with more than two variables, the truth table may become large and cluttered, and students may easily become confused about the order of computation. Though I have noticed this error on several occasions, I could not locate an actual example to reproduce here.

Students seem aware of this confusion and use various strategies to avoid it. One strategy is to use different colours for different columns: maybe pencil for the initial columns (1,4 and 6 in the example above); black pen for intermediate connectives (columns 3 and 5) and red pen for the main connective (column 2). A similar strategy is to highlight the more important columns (that is those which are calculated later) by drawing circles or squares around individual values. Many students will indicate the main connective (ie the final outcome of the truth analysis) by encircling the whole column or at least pointing to it with an arrow. Yet another strategy is to offset the values in each row as exemplified by the student's truth table in Figure 7.13.

P	Q	(P	v	(~	P	⇒	Q))	⇒	(~	Q	⇒	P)
T	T	T			T		T			T		T
				F					F			
						T					T	
			T									
								T				
T	F	T			T		F			F		T
				F					T			
						T					T	
			T									
								T				
F	T	F			F		T			T		F
				T					F			
						T					T	
			T									
								T				
F	F	F			F		F			F		F
				T					T			
						F					F	
			F									
								T				

Figure 7.13 — Use of Offsetting to Highlight Order of Operations

An associated error is for a student to construct correct truth tables, but indicate the wrong column as the main connective. This may have happened, for instance, with a student who constructed correct truth tables for the wffs ~(PvQ) and ~P&~Q, but who then declared that the two wffs were *not* equivalent.

5. T&F: When either conjunct is false, the overall conjunction is also false. But this may be incorrectly judged to be true.

6. Incorrect rows: A truth table for an expression with n variables should have 2^n rows which are normally ordered systematically. But some of these rows may be missing, repeated or the order of rows may be confused.

   For instance, one student constructed a truth table with sixteen rows (half of which were repeated) for a wff with only three variables. Another student made a similar mistake by constructing a truth table with eight rows for a wff with only two variables.

   The ordering of rows is in one sense irrelevant since a truth table may be filled in correctly with the rows in any order. However, if students who do not adhere to a standard order are more likely to leave out a row or to repeat a row. Furthermore, it is difficult to compare truth tables whose rows are not in the same order. In 1994 (when only truth tables were taught) 30% of the group did not use the standard order of rows in Question 3 (see section 7.2.7).

7.4.2 Typical Errors with PMDs

1. Orientation: As discussed above (Section 7.2.1.1), it is important that the implicit labeling of a PMD is agreed upon by all. But students may use an alternate orientation to the convention.

2. Two variables instead of three: If a boolean expression contains three distinct variables, then it is necessary to use a three-variable PMD. But some students do not recognise this and try to use a two-variable PMD. Figure 7.14 shows an example from the 1993 Final Exam, Question 5. This example shows the additional error of forming a disjunction by the match operation instead of the overlay operation, but this is a rare mistake (the only occurrence I have noticed).

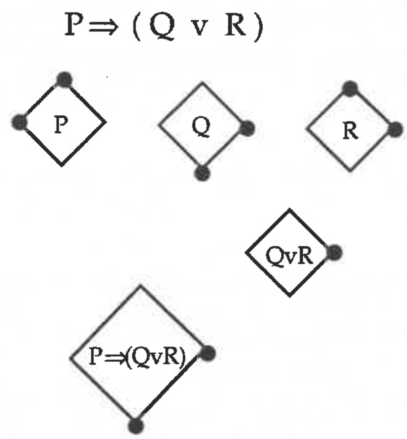


Figure 7.14 — Incorrect PMD for P⇒(QvR)

3. Diagram for Q: In a two-variable PMD, the second propositional variable (say Q) is represented by Figure 7.16a. But some students mistakenly draw the diagram in Figure 7.16b. Two students had previously drawn a correctly labeled two-variable PMD (Figure 7.15), and yet represented Q by Figure 7.16b.

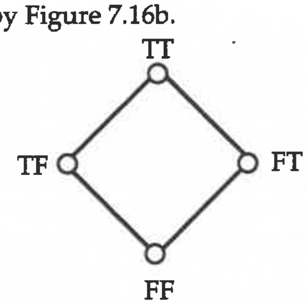


Figure 7.15 — Correctly labeled PMD for two variables

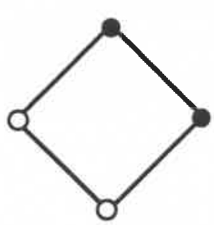


Figure 7.16a — Correct PMD for Q

Figure 7.16b — Incorrect PMD for Q

4. Material implication: When two PMDs are to be combined by the operator "⇒", the correct procedure is to reverse the first diagram and overlay it on the second. But students may not perform this procedure correctly. The fact that this error occurs far less than the corresponding errors with truth tables is one of the strong points in favour of teaching PMDs.

An example from the 1993 Final Exam (Question 5) is given in Figure 7.17. The student has constructed PMDs for two different wffs and concludes incorrectly that they are not equivalent. In both cases the student has constructed the PMDs correctly except for the step where the ⇒ sign is processed. It is probable that the student performed a "reverse and match" operation instead of the correct "reverse-and-overlay". This explanation is consistent with both examples. This student had correctly used reverse-and-overlay for material implication in earlier two-variable questions in the same exam.

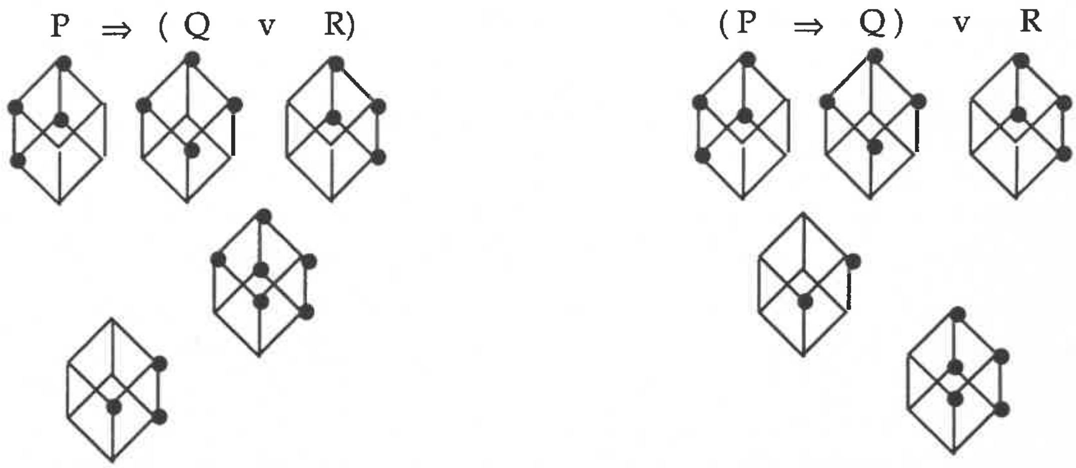


Figure 7.17 — Incorrect use of Material Implication

Another example of a mistake in processing a material implication is shown in Figure 7.18. This example is taken from Question 3 in the 1993 Final Exam. The student has constructed the PMD correctly except for the step where the PMD for ~P is combined with the PMD for Q. In other questions in the same exam this student did use the correct reverse-and-overlay procedure. It is possible that the student was confused by the need to reverse a PMD which was already a reversal of the basic PMD for P (ie a confusion regarding double negation, rather than a confusion regarding material implication).

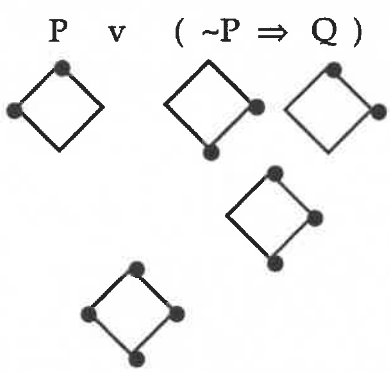


Figure 7.18 — Incorrect use of Material Implication

5. Negation: When a PMD is to be negated, the correct procedure is to *reverse* each corner of the diagram. But some students seem to reflect the diagram instead. For instance, consider the PMD in Figure 7.19 which was constructed for the wff ~(PvQ) by a student.

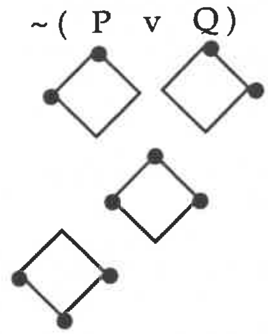


Figure 7.19 — Incorrect use of Negation

Although the first three diagrams are correct, the fourth diagram is not the negation of the third. It seems that the student has mentally either turned the third diagram upside-down or has reflected the diagram about the horizontal axis. This is a reasonably common error when students first encounter negation of PMDs. Perhaps the term "reverse" is ambiguous. Although a number of students have misinterpreted the reverse procedure in this way when they are first taught, the mistake is usually quickly corrected by emphasising that "reverse" means "reverse each corner". The situation in Figure 7.19 is the only case I have seen in which the mistake has persisted.

6. Conjunction: When two PMDs are to be combined by the operator "&", the correct procedure is to *match-the-dots* of the two diagrams. But students may not perform this correctly.

7. Missed a diagram: Constructing a PMD for an expression requires a number of subordinate PMDs to be drawn. But one or more of these may be missed. Related to this is the possible mistake of processing operators in the wrong order, though no example of this has been observed. These two mistakes are most likely to occur when students do not line up the diagrams under the symbols.

In an otherwise very good answer to Question 4 of the 1993 Final Exam, one student constructed the PMDs shown in Figure 7.20. She added notes to the diagrams to indicate exactly which operations were being used and added further emphasis by repeating some PMDs unnecessarily. However, she failed to draw any PMDs for the clause "&(~S⇒~E)". Given this student's overall ability and the quality of her answers to other questions in this exam, I would understand this error as a minor oversight.

This error raises the question of whether the method of processing PMDs promotes such mistakes. The very low frequency of such errors (only three throughout the available data) indicates that this is not the case. This type of error certainly occurs less frequently than the corresponding error with truth tables (number 4 in Section 7.4.1).

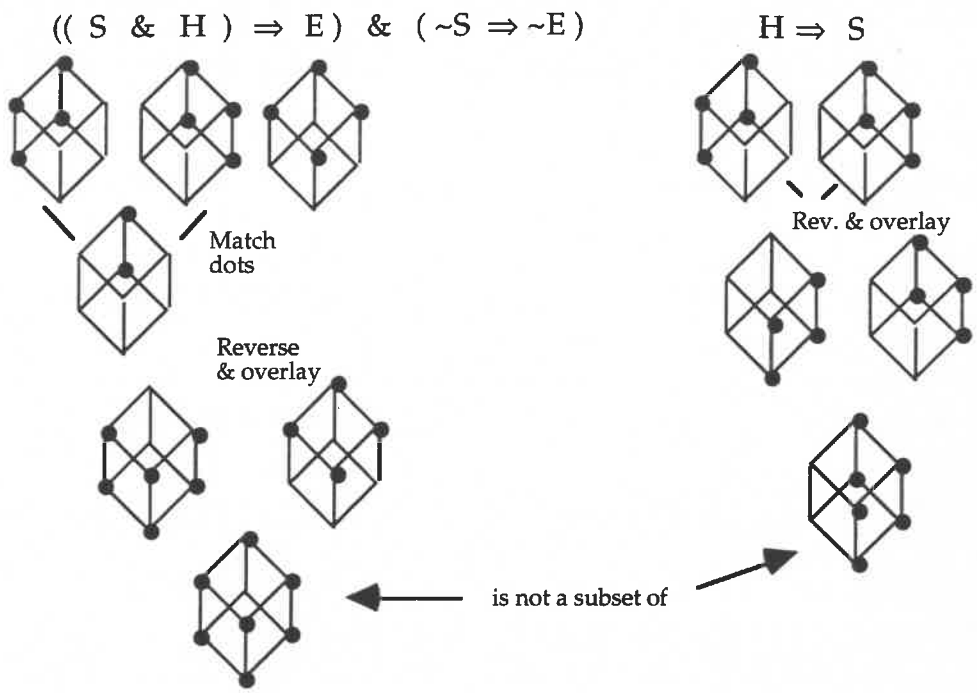


Figure 7.20 — Incorrect — Missing Diagrams

Another example of this error comes from Question 1 in the Assignment of 1993 and is shown in Figure 7.21. In the right half of Figure 7.21 the student has (quite appropriately) left out the diagrams for R, P and Q. However, in the left half, although a diagram has been constructed for Q⇒R, none was constructed for P⇒(Q⇒R), and so the final diagram represents (Q⇒R)⇒(~R⇒(P⇒~Q)) rather than (P⇒(Q⇒R))⇒(~R⇒(P⇒~Q)).

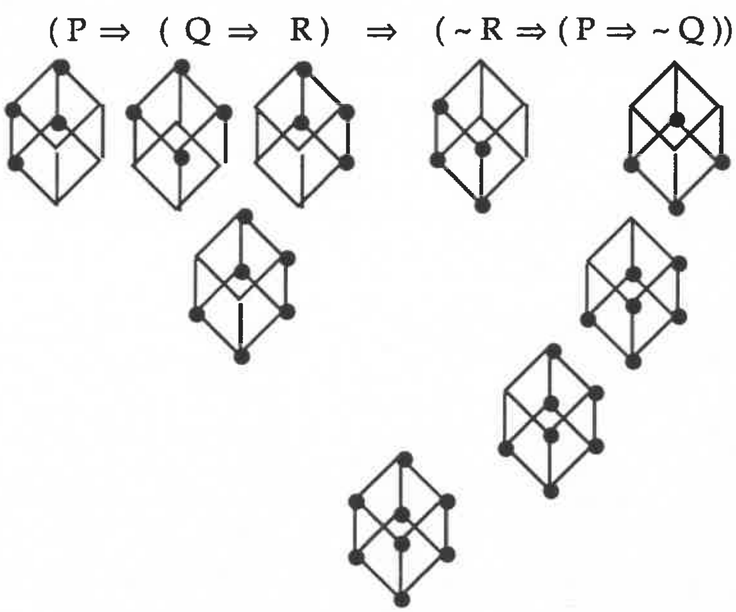


Figure 7.21 — Incorrect — Missing Diagrams

7.5 Student Opinions of the PMD Approach

At the end of the logic courses in both 1992 and 1993, students were requested to complete a course evaluation questionnaire. The University of Natal has no policy on such course and lecturer assessment, and this questionnaire was developed over several years by myself in consultation with other academic staff. The 1992 questionnaire is shown in Figure 7.22. This was administered after the lectures had been completed but before the final exam. The questionnaire contains three sections: The Subject Content, The Lecturer, and Student Information. The 1993 questionnaire differed from the 1992 questionnaire in only two places: Question 7 was slightly modified to read "The logic assignment was a waste of time"; and Question 21 was completely changed to read "Having a weekly quiz followed by discussion of readings promotes learning more than the usual lecturing approach".

Data from these questionnaires was collated in order to give an indication of the students' satisfaction with the course. The following procedure was used —

- For each positively-phrased question, the "average" response was calculated.
- For each negatively-phrased question the responses were first inverted (on the assumption that a response of "5" to a negatively-phrased question can be treated as a response of "1" to a positively-phrased question) and then the average was calculated.
- Careful note was taken of any question which rated significantly less than the others, and goals were set to improve that aspect of the course.
- The overall average for all responses to Questions 1 to 11 was calculated and converted to a percentage to provide an estimate of overall student satisfaction with the Subject Content.
- In a similar way, the overall average for all responses to Questions 13 to 22 was calculated and converted to a percentage to provide an estimate of overall student satisfaction with the Lecturer.
- All written comments (Questions 12 and 23) were noted and recurring comments were closely examined to see whether they indicated areas of the course or details regarding the lecturer which could be improved.

While this procedure provides a very rough guide to student satisfaction, it was considered to be sufficient to assist the lecturer in improving both his lecturing style and the course. It also allows some comparison between courses taught by the same lecturer. Tables 7.12 to 7.15 show the results of those aspects of this evaluation procedure which are relevant to course content.

Course and Lecturer Evaluation
Logic Component of Computer Science 1

Indicate the strength of your agreement or disagreement with the statements below by circling the appropriate digit. Use the following scale —

0 - Inappropriate or do not know
1 - Disagree strongly
2 - Disagree
3 - Neutral
4 - Agree
5 - Agree strongly

The Subject Content

1. The aims of the subject were made clear to me .... 0 - 1 2 3 4 5
2. These aims were achieved .... 0 - 1 2 3 4 5
3. This subject was relevant to my course and career .... 0 - 1 2 3 4 5
4. The method of assessment did not properly test my abilities .... 0 - 1 2 3 4 5
5. I found the subject matter very difficult .... 0 - 1 2 3 4 5
6. The printed notes were useful .... 0 - 1 2 3 4 5
7. The predicate logic assignment was a waste of time .... 0 - 1 2 3 4 5
8. This subject expected too much of me .... 0 - 1 2 3 4 5
9. The tutorial exercises did not teach me anything .... 0 - 1 2 3 4 5
10. The LemmonAid software was very helpful .... 0 - 1 2 3 4 5
11. Overall, I found the course beneficial .... 0 - 1 2 3 4 5

12. Any general comments, or suggestions about the subject matter?

The Lecturer

13. The lecturer spoke clearly .... 0 - 1 2 3 4 5
14. Individual lectures were very disorganised .... 0 - 1 2 3 4 5
15. The lecturer lacked enthusiasm for the subject .... 0 - 1 2 3 4 5
16. The subject matter was developed logically throughout the course .... 0 - 1 2 3 4 5
17. The lecturer responded well to questions during lectures .... 0 - 1 2 3 4 5
18. The lecturer wrote legibly on the chalkboard .... 0 - 1 2 3 4 5
19. The lecturer showed inadequate knowledge of this subject .... 0 - 1 2 3 4 5
20. The lecturer was helpful when I asked for assistance outside class time .... 0 - 1 2 3 4 5
21. Good use was made of visual aids .... 0 - 1 2 3 4 5
22. Overall, the lecturer was very good .... 0 - 1 2 3 4 5

23. Any general comments about the lecturer?

Student Information

24. I attended all the lectures .... 0 - 1 2 3 4 5
25. I attended all the tutorials .... 0 - 1 2 3 4 5
26. On average, I spent the following number of hours each week (outside lectures) on this course ....

Figure 7.22 — 1992 Course Evaluation Questionnaire

Table 7.12 — Comments by Students About the Subject Content (Question 12, 1992)

n=37 Positive Comments	Negative Comments
Interesting (5)[7]	Tutorials were not really beneficial
It was an astounding experience to learn all the different implications that a single sentence could contain	Tutor not familiar enough with the work
Very exciting — especially Predicate Logic	Insufficient time in tutorials. Better to be given the tutorial work beforehand
	Not relevant enough to computer science (5)
	Very intense and involved
	Class tests did not properly test my abilities at constructing proofs
	More practical assessment are needed
	Too difficult (2)
	Make more books available at the library
	More extensive course notes required (2)
	Not enough examples in the course notes
	Heavy workload
	Should leave out Inductive Logic

[7] Where more than one student made the same or very similar responses, the number of such students is shown in brackets following the comment.

Table 7.13 — Comments by Students About the Subject Content (Question 12, 1993)

n=29 Positive Comments	Negative Comments
Helpful and beneficial	Quite theoretical and in a way abstract. We could apply it to everyday life — but we don't — haven't applied it so rigorously. Therefore a bit boring to keep up with. No wonderful new discoveries or exciting concepts.
The notes are useful as they give us a general good outline of the course	The subject was not difficult, but difficult to grasp
Very new approach to the world which made it interesting	Spend more time on proofs, especially the harder aspects (2)
After the first section [Induction], as things became more mathematical it got easier for me	Very confusing and hard to grasp
Enjoyable — I should have got more involved in the discussions	Could not quite make the connection between it and computer science (3), but it helped in my growth analysis experiments
Interesting	More tutorial time would be beneficial
Very good	Discussions may come up with even better improvements. If it can be arranged somehow that every student must have a contribution in these discussions
	The section on Induction is not relevant to computer science, and it was difficult
	Not interesting
	Predicate and Propositional logic difficult — should spent more time on examples as a class rather than in LemmonAid
	During the end of the course I feel practicals should be stopped as students do not have time to study for exams
	Hard to adapt to at first

Table 7.14 — Averages of Responses to Course Evaluation Questionnaire

Question		1992	1993
1	The aims were clear	3.7	3.6
2	The aims were achieved	3.5	3.4
3	Subject was relevant	3.2	3.5
4	Assessment was appropriate	2.5	2.6
5	Subject was very difficult	2.9	2.8
6	Printed notes were useful	3.7	4.7
7	Assignment was a waste of time	2.4	1.9
8	Subject expected too much	3.4	2.0
9	Tutorials didn't teach anything	1.9	2.0
10	LemmonAid was useful	3.7	4.3
11	Beneficial course overall	3.8	3.9

Table 7.15 — Hours Spent on the Logic Course per Week (percentages of student responses)

	1992	1993
Less than 1 hour	37	32
Between 1 and 2 hours	23	25
Between 2 and three hours	17	29
Between 3 and 4 hours	7	0
Over 4 hours	17	14

The most useful results are those instances where the questions were answered differently from one year to the next. During this time I had refined the PMD approach, produced better written notes, introduced the weekly quiz and attempted to improve the course in response to the issues criticised by students in the first questionnaire. Table 7.14 shows five questions whose average response for the two years differ by more than 0.2. In each case the change was positive. The following discussion addresses these seven questions in turn.

Question 3 — On the issue of relevance to course and career, students often fail to see how this course is relevant to computer science. However, I have attempted to include many more examples of logic in computing and this may have lead to the improvement from 3.2 in 1992 to 3.5 in 1993.

Question 6 — The Course Notes given to students in 1992 were completely different to those in 1993. The latter are shown in the Appendix, while the former were very brief and did not describe PMDs at all. As expected, the students in 1993 endorsed the Course Notes to a much higher degree (4.7) than the 1992 students (3.7).

Question 7 — The 1993 Assignment was much broader than the one in 1992, which focused purely on predicate logic. The students seem to have appreciated this change.

Question 8 — In both years, the students generally accepted that the course was pitched at an appropriate level as shown by both Questions 5 ("difficulty") and 8 ("too much was expected"). In my perception, the weekly quiz added to the amount of work because it meant that students had to read and think prior to the first lecture every week. However, the students in 1993 showed an even stronger inclination (4.0) to deny that "This subject expected too much of me" than the 1992 students who did not have weekly quizzes (3.6).

Question 10 — In 1992, the LemmonAid software was described verbally to the students but its use was optional. Based on a positive response by those who did make use of LemmonAid in 1992, several tutorial sessions were devoted to LemmonAid in 1993 and students were encouraged to use the program to assist with the Assignment. The software appears to be a friendly tool which increases the students' confidence with propositional and predicate logic proofs.

The student response to Question 21 in 1993 was also encouraging. This question proposed that "Having a weekly quiz followed by discussion of readings promotes learning more than the usual lecturing approach". I was most interested in the students' responses to this question because I hoped that the new approach (weekly readings from good course notes, followed by a quiz and tutorial-style discussion and exercises on areas which the students identified as unclear) would be well received. The rating of 4.4 indicates quite strongly that students preferred this approach over the more traditional style of lecturing. Not only was the average response high, but only one student disagreed with the statement.

In 1992 the rough metrics described above indicate satisfaction with the course at 72%; in 1993, the figure was 77%. These figures mean very little on their own, but allow useful comparison between this course in logic and other courses I have taught. From similar questionnaires for other courses, I have come to expect satisfaction ratings of between 66 and 78 percent for the course content. I do not want to make any major point by quoting these statistics except that they (along with the student comments listed in Tables 7.12 and 7.13) support my personal feeling that students finish this logic course quite pleased with both the knowledge they have acquired and the means by which that knowledge was imparted.

Chapter 8: Conclusion

8.1 Technical Conclusions

This dissertation has presented the use of Possible Models Diagrams as a tool for analysing formulae and sequents of propositional logic. A PMD for *n* variables is graph G which is isomorphic to the hypercube Q_n, whose vertices are n-tuples indicating the assignment of truth values to the *n* variables. The vertex set is partitioned into the two sets T(G) (the darkened vertices) and F(G) (the vertices left open).

Every propositional wff can be expressed as an induced subgraph of a hypercube [HARA89] and one way to represent this induced graph is in the form of a PMD. Every propositional wff can be expressed as a PMD in which the set T(G) indicates those possible models for which the wff turns out to be true. Compound PMDs may be constructed from simpler PMDs using the operations reverse, overlay and match and there exist simple algorithms for applying these operations either by hand or by computer.

Once a PMD has been constructed for a wff, it is easy to determine whether the wff is tautologous, contingent or inconsistent. Informally, one only has to see whether all the nodes of the PMD are darkened, whether only some are darkened or whether none are darkened respectively. More formally, if the PMD G is such that F(G) is empty, then the corresponding wff is tautologous; if T(G) is empty then the wff is inconsistent; if neither T(G) nor F(G) are empty, then the wff is contingent.

Apart from analysing individual wffs, PMDs may be used to analyse sequents. A propositional sequent $A_1, A_2, \ldots, A_n \therefore C$ is valid if and only if the PMD for the conjunction of the premises (say L) and the PMD for the conclusion (say R) are so related that $T(L) \subseteq T(R)$.

PMDs have the same expressive power as truth tables, that is, they are informationally equivalent in Simon's sense. It is not possible to generalise the PMD approach (nor truth tables) to include predicates and quantifiers.

Whereas this research has examined the use of PMDs for teaching basic propositional logic, this is by no means the only area in which they can be applied. Gupta and Pratt have use the same fundamental idea as the basis for defining concurrent processing [GUPT93]. Brink and Heidema have used the partial ordering implicit in PMDs as a way of analysing the verisimilitude of theories [BRIN87].

8.2 Pedagogical Conclusions

Although PMDs and truth tables are informationally equivalent, they are not computationally equivalent (in Simon's sense). When used for human problem solving, certain information can be more readily retrieved from a PMD than from a truth table. This is because they are more visual and imagistic than truth tables, because labeling is implicit, and because they are iconic rather than purely symbolic (in Bruner's sense).

Although PMDs are defined in terms of graphs and boolean algebra, it is easier to introduce PMDs to students using the basic concepts of set theory. Student's pre-existing knowledge of set concepts such as complementation, union, intersection and subset can be used as the basis for teaching the logical concepts of negation, disjunction, conjunction and material implication respectively. Linking disjunction to set intersection settles the question of why disjunction is defined to be inclusive rather than exclusive. Linking material implication to the subset relation avoids the typical student opposition to a truth-functional definition of material implication.

Three simple visual operations allow a PMD to be built for any propositional wff — the action of *overlaying* two PMDs corresponds to forming a disjunction, *matching* the dots of two PMDs corresponds to forming a conjunction, and *reversing* the dots of a PMD corresponds to negation. Material implication is acheived by a combination of the reverse and overlay operations.

With all forms of symbolic logic, including the PMD approach, there is a threat that mechanical learning may overshadow concept learning. In order to avoid this threat, students should be presented with a multiplicity of logical representations and should become experienced in translating between those representations. In this way students are encouraged to learn not just the mechanical processes required by any one representational system, but also to learn the abstract concepts underlying those representations and processes. Given the body of evidence indicating the difficulty of transferring formal concepts to practical applications, it is especially important that students be given exercises requiring translation between informal English descriptions of logical problems and symbolic notation.

A course on logic for first-year computer science students has been designed and taught for several years based on these ideas. Various forms of logic (informal, inductive, propositional, predicate, fuzzy, probabilistic, multi-valued and modal) and various logical tools (truth tables, PMDs, and a natural deduction system) are all presented in this course. PMDs have been found to fit well in such a course.

The available quantitative data drawn from assessment of students over three years suggests that when PMDs are used the odds of giving correct answers is greater than the corresponding odds when truth tables are used. There is a strong tendency among students to use PMDs rather than truth tables for situations involving two distinct propositional variables when given the choice, although this preference is reduced when the number of propositional variables increases to three.

In addition, a variety of qualitative data suggests that student difficulties with truth tables are alleviated by the use of PMDs. In particular, it has been noted that material implication is a difficult concept to master and that when students use truth tables the majority of their mistakes involve material implication. However, relatively few errors involving material implication are made when students use PMDs. Furthermore, a number of students draw the PMD icon for material implication to help them fill in truth tables correctly.

In course assessment surveys, students evaluate the PMD approach positively.

8.3 Future Research Directions

The topic of transfer has been raised a number of times throughout this dissertation. One would hope that a course in formal logic equips students to apply logical principles in less formal situations. However, the extent to which this occurs, if it occurs at all, is still an open question. Further research with PMDs is likely to focus on whether they assist in this task of transfer.

A study has been commenced in which students are asked to judge the veracity of conjunctions, disjunctions and conditional statements (all presented in English), given certain assurances about the veracity of the components of those statements. The statements are designed to enable the students' internal "truth table" to be inferred from their answers. The same questions will be asked at the commencement of the course and at its completion in order to ascertain whether the students' internal "truth tables" are altered by the course.

The changing shape of South African culture and scholarship also provides fertile ground for cross-cultural studies. Are there cultural differences in the perception of logic? The bi-lingual studies which Zepp has undertaken with Cantonese, Sesotho and English speakers [ZEPP82 and ZEPP87] could be extended to include Zulu speakers. Would the visual nature of PMDs assist black African students to learn symbolic logic?

Bibliography

Since portions of this dissertation have been concerned with historic trends in the teaching of logic, it has been important to record the initial dates of publication. Thus a bibliographical citation such as [LEMM65, p57] refers to the book by Lemmon which was first published in 1965, although the actual text available to this author was the 1971 edition. This Bibliography lists the editions to which the cited page number (p57 in this case) refers.

[AHO92] Aho,A.V. and Ullman,J.D.; Foundations of Computer Science, Freman, 1992

[ALLE92] Allen,C. & Hand,M.; Logic Primer, MIT, 1992

[AMBR48] Ambrose,A. and Lazerowitz,M.; Fundamentals of Symbolic Logic, Rinehart, 1948

[ASL94] Association of Symbolic Logic; A.S.L. Guidelines for Logic Education, the Report of the Committee on Logic Education of the ASL, chaired by Jon Barwise; 1994

[BART60] Bartee,T.C.; Digital Computer Fundamentals, McGraw-Hill, 1977

[BARW93] Barwise,J.; Heterogeneous Reasoning, in *Working papers on diagrams and logic*, ed. Barwise,J. and Allwein,G., preprint No. IULG-93-24, Indiana University Logic Group, 1993

[BASS53] Basson,A.H. and O'Connor,D.J.; Introduction to Symbolic Logic, University Tutorial, 1970

[BELL77] Bell,J.L. and Machover,M.; A Course in Mathematical Logic, North-Holland, 1977

[BETH59] Beth,E.W.; The Foundations of Mathematics, North Holland, 1959

[BOOL1854] Boole,G.; An Investigation of the Laws of Thought, Dover 1958 (first published by Macmillan in 1854)

[BRAD79] Bradley,R. and Swartz,N.; Possible Worlds, Blackwell, 1979

[BRIN87] Brink,C. and Heidema,J.; A verisimilar ordering of theories phrased in propositional language, in *British Journal for the Philosophy of Science*, vol 38, 1987, pp533-549

[BROW75] Brown,J.S. and Burton,R.R; Multiple representations of knowledge for tutorial reasoning in *Representation and Understanding*; ed. Bobrow,D.G and Collins,A; Academic, 1975

[BRUN67] Bruner,J.; Toward a Theory of Instruction, Harvard University, 1967

[CARN64] Carney,J.D. and Scheer,R.K.; Fundamentals of Logic, Macmillan, 1980

[CARR1896] Carroll,L.; Symbolic Logic, Harvester, 1977 (Part 1 was first published in 1896)

[CHEN85] Cheng,P.W. andHolyoak,R.; Pragmatic Reasoning Schema in *Cognitive Psychology*, vol 17, 1985, pp391-416

[CHEN86] Cheng,P.W. Oliver,L. and Nisbett,R.E.; Pragmatic versus syntactic approaches to training deductive reasoning in *Cognitive Psychology*, vol 18, 1986, pp293-328

[CHUR44] Church,A.; Introduction to Mathematical Logic (vol 1), Princeton, 1956

[CLAR93] Clarke,M.C.; Possible Models Diagrams: a visual alternative to truth tables in the ACM SIGCSE Bulletin, v25, #1, March 1993

[CLAR94] Clarke,M.C.; Visualising boolean operations on a hypercube, to appear in *Mathematical and Computer Modelling*

[COLL85] Collins,A., Brown,J.S. and Newman,S.E.; Cognitive Apprenticeship: teaching the crafts of reading, writing and mathematics in *Knowing, Learning and Instruction: Essays in honor of Robert Glaser*, ed. Resnick,L.B., LEA, 1989

[COOL42] Cooley,J.C.; A Primer of Formal Logic, Macmillan, 1942

[COPI61] Copi,I.M.; Introduction to Logic, Macmillan, 1967

[COX93] Cox,R. and Oberlander,J.; Graphical effects in learning logic: reasoning, representation and individual differences in *Semantic Issues in Graphical Representations*, ed. Oberlander,J., ESPRIT Basic Research Action P6296, Deliverable 3.1, 1993

[CREI1898] Creighton,J.E.; An Introductory Logic, Macmillan 1904

[DAVI86] Davis,W.A.; An Introduction to Logic, Prentice-Hall, 1986

[DENI89] Denis,M.; Image and Cognition, Simon and Schuster, 1991

[DOWS86] Dowsing,R.D., Rayward-Smith,V.J. and Walter,C.D.; A First Course in Formal Logic and its Applications in Computer Science, Blackwell, 1986

[EATO31] Eaton,R.M.; General Logic: an introductory survey, Charles Scribner, 1931

[EISI93] Eisinger,N. and Ohlbach,H.J.; Deduction Systems Based on Resolution, in *Handbook of Logic in Artificial Intelligence and Logic Programming*, ed. Gabbay,D.M., Hogger,C.J. and Robinson,J.A., Oxford Uni, 1993

[EVAN82] Evans,J.St.B.T.; The Psychology of Deductive Reasoning, Routledge and Kegan Paul, 1982

[FISH88] Fisher,A.; The Logic of Real Arguments, Cambridge University, 1988

[FITC52] Fitch,F.B.; Symbolic Logic: an introduction, Ronald, 1952

[FOWL1895] Fowler, T.; Logic: Deductive and Inductive, Clarendon, 1895

[GALT90] Galton,A.; Logic for Information Technology, Wiley, 1990

[GARD85] Gardner,H.; The Mind's New Science; Basic, 1987

[GAVE85] Gavelek,J.R. and Raphael,T.E.; Metacognition, Instruction, and the Role of Questioning Activities in *Metacognition, Cognition, and Human Performance* (vol 2); ed. Forrest-Pressley,D.L., MacKinnon,G.E. and Waller,T.G; Academic Press, 1985

[GEAC76] Geach,P.T.; Reason and Argument, Blackwell, 1976

[GENT35] Gentzen,G.; The Collected Papers of Gerhard Gentzen, ed Szabo,M.E, North-Holland, 1969

[GEOG79] Georgacarakos,G.N. and Smith,R.; Elementary Formal Logic, McGraw-Hill, 1979

[GILH88] Gilhooly,K.J.; Thinking: directed, undirected and creative, Academic, 1988

[GRIE93] Gries,D. and Schneider,F.B.; A Logical Approach to Discrete Math, Springer Verlag, 1993

[GROS90] Grossen,B. and Carnine,D.; Diagramming a logic strategy: effects on difficult problem types and transfer in *Learning Disability Quarterly*, vol 13, 1990, pp168-182

[GUPT93] Gupta,V. and Pratt,V.R.; Gates Accept Concurrent Behavior, in *Proc. 34th Ann. IEEE Symp. on Foundations of Computer Science*, 1993, pp62-71

[GUTT71] Guttenplan,S.D. and Tammy,M.; Logic: a comprehensive introduction, Basic, 1971

[HAMI78] Hamilton,A.G.; Logic for Mathematicians, Cambridge University, 1978

[HARA88] Harary,F., Hayes,J.P. and Wu,H-J; A survey of the theory of hypercube graphs in *Computers and Maths with Applications*, v15, #4, 1988.

[HARA89] Harary,F.; The graph of a boolean function, in the *Journal of Experimental and Theoretical Artificial Intelligence*, #2, 1989

[HART74] Hartshorne,C. and Weiss,P. (ed.); Collected Papers of Charles Sander Peirce (6 volumes), Harvard University, 1974

[HERM73] Hermes,H.; Introduction to Mathematical Logic, Springer, 1973

[HILB28] Hilbert,D. and Ackermann,W.; Principles of Mathematical Logic, Chelsea 1950

[HOCU79] Hocutt,M.; The Elements of Logic Analysis and Inference, Winthrop, 1979

[JEFF67] Jeffrey,R.C.; Formal Logic: its scope and limits, McGraw-Hill, 1967

[JOSE06] Joseph,H.W.B.; An Introduction to Logic, Clarendon 1906

[JOYC56] Joyce,G.H.; Principles of Logic, Longmans, Green & Co, 1956

[KELL90] Kelley,D.; The Art of Reasoning, Norton, 1990

[KEYN1884] Keynes,J.N.; Studies and Exercises in Formal Logic, Macmillan, 1894

[KIRW78] Kirwan,C.; Logic and Argument, Duckworth, 1978

[KNEA62] Kneale,W. and Kneale,M.; The Development of Logic, Clarendon, 1962

[KNEE63] Kneebone,G.T.; Mathematical Logic and the Foundations of Mathematics, Van Nostrand, 1963

[KORF74] Korfhage,R.R.; Discrete Computational Structures, Academic, 1974

[LARK87] Larkin,J.H. and Simon,H.A.; Why a diagram is (sometimes) worth ten thousand words in *Cognitive Science*, vol 11, 1987, pp65-99

[LEMM65] Lemmon,E.J.; Beginning Logic, Nelson's University Paperbacks, 1971

[LU89] Lu, Z; Mathematical Logic for Computer Science, Word Scientific, 1989

[MANN85] Manna,Z. and Waldinger,R.; The Logical Basis for Computer programming, Addison-Wesley, 1985

[MARK86] Markovits,H.; Drawings and conditional reasoning in *Educational Studies in Mathematics*, vol 17, 1986, pp81-87

[MASS70] Massey,G.J.; Understanding Symbolic Logic, Harper and Row, 1970

[MATE65] Mates,B.; Elementary Logic, Oxford University, 1972

[MCCA81] McCawley,J.D.; Everything linguists have always wanted to know about logic but were ashmed to ask, Blackwell, 1981

[MELL04] Mellone,S.H.; An Introductory Text-Book of Logic, Blackwood, 1937

[MEND64] Mendelson,E.; Introduction to Mathematical Logic, Van Nostrand, 1964

[MITC62] Mitchell,D.; An introduction to Logic, Hutchinson University, 1962

[NEWS83] Newstead,S.E. and Griggs,R.A.; The language and thought of disjunction in *Thinking and Reasoning: psychological approaches*, ed. Evans,J., Routledge and Kegan Paul, 1983

[NICK85] Nickerson,R.S., Perkins,D.N. and Smith,E.E; The Teaching of Thinking, Lawrence Erlbaum, 1985

[NISB87] Nisbett,R.E, Fong,G.T., Lehman,D.R and Cheng,P.W.; Teaching Reasoning in *Science*, vol 238, 1987, pp625-631

[OBER94] Oberlander,J., Cox,R. and Stenning,K.; Proof Styles in Multimodal Reasoning, presented at the 4th International Conference on Information-oriented Approaches to Language, Logic and Computation, 1994

[OBRI72] O'Brien,T.C.; Logical thinking in adolescents in *Educational Studies in Mathematics*, vol 4, 1972, pp401-428

[OBRI73] O'Brien,T.C.; Logical thinking in college students in *Educational Studies in Mathematics*, vol 5, 1973, pp71-79

[PAUL87] Paulson,L.C.; Logic and Computation, Cambridge University, 1987

[PETE94] Peterson,D.M. (ed); Forms of Representation, Intellect, 1994

[POSP74] Pospesel,H.; Introduction to Logic: Propositional Logic, Prentice-Hall, 1984

[POSP76] Pospesel,H.; Introduction to Logic: Predicate Logic, Prentice-Hall, 1976

[POSP78] Pospesel,H. and Marans,D.; Arguments: deductive logic exercises, Prentice-Hall, 1978

[PRIO55] Prior,A.N.; Formal Logic, Clarendon, 1955

[QUIN40] Quine,W.V.O.; Mathematical Logic, Norton, 1940

[QUIN41] Quine,W.V.O.; Elementary Logic, Ginn, 1941

[QUIN50] Quine,W.V.; Methods of Logic, Holt,Rinehart and Winston, 1982

[READ1898] Read,C.; Logic: Deductive and Inductive, Alexander Moring, 1914

[REEV90] Reeves,S. and Clarke,M.; Logic for Computer Science, Addison-Wesley, 1990

[REIC47] Reichenbach,H.; Elements of Symbolic Logic, Macmillan, 1947

[ROBI79] Robinson,J.A.; Logic: form and function, Edinburg University, 1979

[ROSS53] Rosser,J.B.; Logic for Mathematicians, McGraw-Hill, 1953

[RUSS06] Russell,J.E.; An Elementary Logic, Macmillan 1906

[RUSS14] Russell,L.J.; An Introduction to Logic (from the standpoint of education), Macmillan, 1914

[SAIN91] Sainsbury,M.; Logical Forms, Blackwell, 1991

[SEID89] Seidman,R.H.; Computer programming and logical reasoning: unintended cognitive effects in the *Journal of Educational Technology Systems*, vol 18(2) 1989-90, pp123-141

[SHAP73] Shapiro,B.J. and O'Brien,T.C.; Quasi-child logics in *Educational Studies in Mathematics*, vol 5, 1973, pp181-184

[SHOE67] Shoenfield,J.R.; Mathematical Logic, Addison-Wesley, 1967

[SIGW1873] Sigwart,C; Logic (two volumes), Swan Sonnenschein, 1895

[SMUL68] Smullyan,R.M.; First-Order Logic, Springer, 1968

[STEB30] Stebbing,L.S.; A Modern Introduction to Logic, Methuen, 1946

[STEB43] Stebbing,L.S.; A Modern Elementary Logic, Methuen, 1952

[STEN92] Stenning,K. and Oberlander,J.; A cognitive theory of graphical and linguistic reasoning: logic and implementation; Research Report HCRC/RP-20, Human Communication Research Centre, University of Edinburgh, 1992

[SUPP57] Suppes,P.; Introduction to Logic, Van Nostrand, 1957

[THEW83] Thewlis,P.J. and Foxon,B.N.T.; From Logic to Computers, Blackwell, 1983

[UTZI82] Utzinger,J.; Logic for Everyone - alternate techniques for teaching logic to learning diasbled students in the university, Washington University, 1982

[VENN1881] Venn,J.; Symbolic Logic, Macmillan, 1894

[VYGO35] Vygotsky,L.S.; Interaction between learning and development in *Mind in Society*, ed. Cole,M., John-Steiner,V., Scribner,S. and Souberman,E., Harvard Uni, 1978

[WASO83] Wason,P.C.; Realism and rationality in the selection task in *Thinking and Reasoning: psychological approaches*, ed. Evans,J., Routledge and Kegan Paul, 1983

[WELT11] Welton, J. and Monahan,A.J.; An Intermediate Logic, University Tutorial, 1923

[WELT20] Welton,J.; Groundwork of Logic, University Tutorial, 1920

[WELT1896] Welton, J.; Manual of Logic (vol 2), University Correspondence College, 1896

[WERK48] Werkmeister,W.H.; An Introduction to Critical Thinking, Johnsen, 1948

[WHIT10] Whitehead,A.N and Russell,B; Principia Mathematica; Cambridge University, 1925

[WITT74] Wittrock,M.C.; A generative model of mathematics learning in *Journal for Research in Mathematics Education*, vol 5, 1974, pp181-196

[ZEPP82] Zepp,R.; Bilinguals' understanding of logical connectives in English and Lesotho in *Educational Studies in Mathematics*, vol 13, 1982, pp205-221

[ZEPP87] Zepp,R., Monin,J. & Lei,C.L.; Common logical errors in English and Chinese in *Educational Studies in Mathematics*, vol 18, 1987, pp1-17

Contents

Preface - Why study logic?

Logic is the study of right reasoning. It is the process of putting together reasons for believing something. Colloquially we use terms such as "logical", "rational", "reasonable" etc very loosely, but in these notes we will establish an exact formal basis for logic.

There are a variety of reasons for believing things, including:

- being told by parents or some authority figure ("medical authorities warn that smoking is a health hazard")
- peer pressure, fear, compulsion
- ignorance
- personal observation ("I <u>saw</u> the man kill her")
- reasoning about observations ("all people die")
- reasoning about other beliefs ("I believe that 5+3=8, and so I can reason that I should also believe that 8-3=5")

Only the last two of these fall within the scope of logic, however, there is not just one way of doing logic. Indeed there are a wide variety of logics - from informal logic whose general principles apply to everyday reasoning, to rigidly structured logic relevant particularly to computer science. Right reasoning is a basic foundation for all academic pursuits, maybe especially for science, and maybe even more especially for computer science. The computer is a machine which rigidly adheres to highly structured logic and those who wish to understand computers must be able to think in the same rigidly logical way.

These notes were written primarily as part of the Computer Science 1 syllabus at Natal University (Pietermaritzburg). They serve as an introduction to the whole field of logic, not just "computer logic" - an introduction which should serve the student throughout their academic endeavours and hopefully the rest of their lives. Nevertheless, the emphasis of these notes is on the kind of formalism most suited to computer science: boolean operations, propositional and predicate calculi, and various forms of numerically-based logic.

Section 1 - Some Basic Concepts

Arguments

eg Thembi expects her income to rise by 10% next year but then realises that prices of goods and services will also probably increase by about 10%. She also knows that having a higher income may put her in a higher tax bracket and so infers that next year she will actually be worse financially off than this year.

Thembi has started with some premises and has derived a conclusion - this process is call an inference. We can describe this process in various ways:

- a conclusion is inferred from some premises
- a conclusion follows from the premises
- the premises imply a conclusion
- the premises lead to a conclusion
- the premises entail a conclusion

In logic, an argument is an inference (or a series of inferences) which has been written down in a structured form. (This is much different from the use of the word "argument" in normal English.) An argument is an ordered list of logically connected reasons for believing something. An argument is a proposed proof of some conclusion.

There are many ways to try to convince someone to believe something which are not technically arguments. For instance, "Drink Extra-Lite Beer! It's the beer of champions!" is a form of persuasion, but not an "ordered list of logically connected reasons for believing something."

Propositions

Each statement in an argument represents a proposition. The argument below contains three propositions -

1. Whenever it rains the grass grows
2. It is raining heavily today
3. Therefore, the grass must be growing

One proposition may be expressed in different ways: for instance, the statement "If it rains then the grass will grow" represents the same proposition as statement 1. above. Similarly, several statements in different languages may all express the same proposition.

Propositions are claims about certain facts. Propositions may be affirmed or denied, which makes them different from questions, commands and exclamations. eg "How are you?", "Pass the salt" and "Hallelujah!" are not propositions.

Both premises and conclusions are propositions. In fact, one proposition may be a conclusion in one context and a premise in another.

Exercise Which propositions are premises and which are conclusions in the following argument?

1. In a legal Pascal program, each BEGIN must have a matching END
2. This program has 5 BEGINs and 4 ENDs
3. This is not a legal Pascal program
4. Illegal programs will not compile successfully
5. This program will not compile

"Truth"

The words "true" and "false" are also used very loosely, and will need to tightened up in order to reason logically. Every proposition can either be true or false: these are the only possible truth values and there is no possibility that any one proposition can be both true and false. This is a basic presumption we will make for most of these notes, and it is best expressed as two Laws -

> Definition: Laws of Rational Thought
> 1. The Law of Non-Contradiction - no proposition can be both true and false simultaneously.
> 2. The Law of Excluded Middle - a proposition must be either true or false, there is no middle ground.

Two justifications can be distinguished for a proposition's truth value -

1. Empirical truth - a proposition is true if it accurately describes reality. This is the realm of science; of observation. The proposition "It is raining" is true if you go outside and discover that it is actually raining, otherwise it is false.
2. Necessary (or logical) truth - the truth or falsity of some propositions can be determined without any need for observational data. eg "Either it is raining or it is not raining" is always true.

Although propositions are either true or false, it doesn't make sense to say that an argument is true or false. Rather, we are concerned about whether the conclusion really does follow from the premises. If so, we say that the argument is valid, otherwise it is invalid.

> Definition: Validity
> An argument is valid whenever the propositions are related in such a way that it is absolutely impossible for the premises to be true unless the conclusion is also true. If it is possible for the premises to all be true but the conclusion false, then the argument is invalid.

Note that the validity of an argument does not require that the premises be true. For instance the raining/grass growing example above is valid, even if it is not raining. A valid argument may proceed from premises which are false and lead to a conclusion which is false! Clearly we need a further way of classifying arguments ...

> Definition: Soundness
> An argument is sound whenever it is valid and all the premises are true. Otherwise it is unsound.

Fallacies

Although these notes are about principles of right reasoning, it is useful to start with an examination of bad reasoning. Fallacies, or invalid forms of reasoning, fall into three categories: psychological, material and logical. Typical examples of these (certainly not a complete list) are described below.

1. Psychological fallacies

- mislead the audience by taking advantage of psychological factors.

Ad hominem

A personal attack which seeks to discredit the source of an argument rather than address the argument itself. May include name calling and character assassination. The attack may or may not be accurate, but nevertheless is irrelevant to the validity or invalidity of the argument.

Appeal to the supporters of the argument

"Scientists all agree that ...", "Everybody knows that ...", "Intelligent people have always believed that ...".

Argument from ignorance

From the inability to prove something, it is not valid to infer that the opposite must be true. eg "There is no proof for the existence of God; therefore he does not exist."

Appeal to emotions

The use of threats to induce compliance by fear; the use of emotive words; an appeal to sympathy (eg in a court room).

2. Material fallacies

- occur when conclusions are based on faulty material (ie premises)

Inconsistent premises

Given a set of premises, there may be a valid argument which leads to some conclusion, yet the argument is unsound because there is something wrong with the premises. Not only might one premise be false, but they might be incompatible (inconsistent) with each other.

> Exercise: What is wrong with the following argument?
> Bigdorp is 20km north of Littledorp and Littledorp is 20km west of Bergville. There is a straight road that starts at Bigdorp and goes through Bergville to Dorpville. Dorpville is north-east of Bergville. Therefore, Littledorp is closer to Bergville than to Dorpville.

Black-or-white fallacy

Assuming that there are only two opposing choices when there are actually other possibilities. eg "Either a person is good or they are bad", "Either we increase our sales or we will go bankrupt", "Either you win or you lose".

Hasty generalisation

Making a general principle from limited experience (eg picking one white ball out of a hat and deducing that they all must be white), or from unrepresentative data (eg carrying out some research in an old-people's home and inferring that all humans have difficulty hearing).

Post-hoc reasoning

Given that event A occurs immediately before event B, it is fallacious to conclude that A caused B. eg Themba had a cold but it went away the day after he started taking Sniffleless Medicine. Nevertheless, he would not be justified in concluding that taking Sniffleless cured the cold. (Because of Hasty Generalisation, he would be even less justified to conclude that "Sniffleless cures colds".)

Begging the question (petitio principii)

This occurs when the proposition to be proved is somehow assumed to start with.

eg Thembi - "That man is insane!"
Themba - "How do you know?"
Thembi - "Look, I can prove it - he's a complete lunatic, therefore he must be insane!"

3. Logical fallacies

- misusing the rules of inference

Definition: Conditional expressions
Propositions of the form "if X then Y" are called conditionals. X is called the antecedent and Y is called the consequent.

Affirming the consequent

The following (invalid) argument affirms the consequent and falsely infers the antecedent -

If a word is a Zulu verb then it ends with an 'a'
The word 'ilanga' ends in 'a'
Therefore 'ilanga' is a Zulu verb

Denying the antecedent

The following (invalid) argument denies the antecedent and falsely infers the consequent -

If it rains then the grass will grow
It is not raining
Therefore the grass is not growing

<u>Exercises</u>

1. For each of the following arguments,

   i) Write down the premises and and the conclusion.
   ii) Write down any unstated premises or conclusions.
   iii) Is the argument sound? Is it valid?
   iv) Name any fallacy you can see.

1.1 All unicorns have four feet. Eugene is a unicorn. Therefore Eugene has four legs.

1.2 The ratio of the circumference of a circle to its diameter is 22/7. Hence any circle with a diameter of 3.5cm will have a circumference of 11cm.

1.3 Any gork that you can show me is guaranteed to have a snozzle. This minjent is obviously not a gork since it isn't snozzled.

1.4 By definition, a Communist is committed to the redistribution of wealth. Now Chris Hani has clearly stated his belief that land and money needs to be taken from the rich and given to the poor and so he must be a Communist.

1.5 All medical researchers agree that smoking increases the chance of contracting lung cancer. The only people who disagree are the tobacco companies and it would be foolish to believe them!

1.6 If Charles Babbage invented the QuickSort algorithm, he would be a famous computer scientist. Charles Babbage <u>is</u> a famous computer scientist. Therefore he must have invented QuickSort.

1.7 Don't touch that!!! It's hot!

Exercises (continued)

1.8 All the evidence goes to show that what we regard as our mental life is bound up with brain structures and organised bodily energy. Therefore it is rational to suppose that mental life ceases when bodily life ceases. (Bertrand Russell)

1.9 If God exists then the grounding reality itself is not ultimately groundless. Why? Because God is then the primal ground of all reality. (Hans Küng)

1.10 As for Federal supervision of elections in the South, I whole-heartedly support the Administration's referee proposal. I believe it is far superior to the well-intentioned but less effective recommendation of the Civil Rights Commission, because the referee proposal will deal not only with registrations, but also with voting itself. After all, what good does it do to be able to register if you cannot vote? (Richard Nixon)

1.11 All governments restrain and rule people; therefore, all governments are totalitarian and authoritarian to some extent or another. (Totalitarianism refers to a centralized form of government in which those in control grant neither recognition nor tolerance to parties differing in opinion. Authoritarianism is a system of governing that calls for unquestioning submission to authority.) (Robert Ringer)

2. Benno Torelli, genial host at Hamtramck's most exclusive nightclub was shot and killed by a racketeer gang because he fell behind in his protection payments. After considerable effort on the part of the police, five men were brought before the Court where they were asked what they had to say for themselves. Each of the men made three statements: two true and one false -

Lefty: "I didn't do it I never owned a revolver. Spike did it."
Red: "I did not kill Torelli. I never owned a revolver. The other guys are all passing the buck."
Dopey: "I am innocent. I never saw Butch before. Spike is guilty."
Spike: "I am innocent. Butch is the guilty man. Lefty lied when he said I did it."
Butch: "I did not kill Torelli. Red is the guilty man. Dopey and I are old pals."

On the basis of these statements, who did it?

3. In a certain area of the deepest, darkest jungle in the Amazon there are two tribes. Members of one tribe invariably tell the truth, while the other tribe always lies. A certain tourist became lost in the jungle and, near to death, comes to a fork in the track. At the fork there happens to be a native but the tourist does not know which she is from.

   i) If the tourist could ask two questions, how could he find out from the native which path to take to civilisation? (This is easy!)
   ii) Unfortunately, the tourist is so weak he only has enough strength to ask one question. Can you think of a single question such that after the native replies, the tourist will know which of the two paths to take?

Section 2 - Induction

Induction is the process by which general conclusions are infered from particular instances. To induce something is to form a general principle based on limited experience. Most inductive arguments rely on four principles: generalisation, causal reasoning, hypothesis formation and refutation, and reasoning by analogy. These are the ways in which most of our everyday learning and reasoning works. We could not survive without making such inferences, however, we will see that all forms of induction are technically invalid.

1. Generalisation

In order to make sense of the world we notice similarities and differences between objects and events and we use these similarities and differences to define categories.

eg1 From my experience that every dog I've met barks, I conclude that all dogs bark.
eg2 The water we've tested boils at 100°C, so we conclude that all water boils at 100°C.
eg3 The sun has risen on every day I can remember, so it will always rise each day.

Basic Principle of Generalisation
Some X is Y
∴ All X is Y

But what about - Some students pass exams
∴ All students pass exams

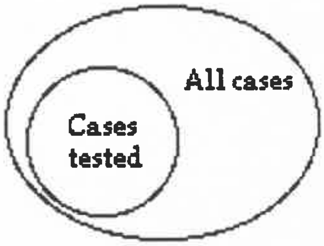
In an inference based on Generalisation, it is possible for the premise to be true but the conclusion false, and hence this form of reasoning is invalid. We could change to principle to make it valid -

Extended Principle of Generalisation
Some X is Y
No X is not Y
∴ All X is Y

...but normally it is not possible to establish the second premise. Nevertheless, it is clear that some generalisations are more reliable than others; given that a generalisation cannot be 100% guaranteed, how can we judge its reliability in order to avoid the fallacy of Hasty Generalisation?

i) Were a sufficient number of samples tested?

ii) Were the samples representative of the whole population?



i) What constitutes "a sufficient number of samples"?
- Depends on your prior knowledge of the situation. For instance you would probably only need one sample to determine that "This whole bucket of milk is sour" but a large number of samples to determine that "All police cars are yellow".
- If all cases in the population are very similar, a small sample size is acceptable, but if there is a lot of variation in the population then many samples should be tested.

- Statistical analysis can make allowance for sample size to some extent but in general the question "how many observations are needed in order to make a reliable generalisation?" is unanswered.

ii) What constitutes "a representative sample"?
- If the population is diverse, choose samples which mirror the diversity.
  - eg_1 When selecting balls from a hat, shake the hat and choose balls from different places in the hat
  - eg_2 When researching whether police cars are yellow, examine police cars from different countries, road patrols, riot squads, drug prosecution etc
- Look for counter-examples.
- Again, in general the question remains unanswered.

Lastly, make generalisations which are as specific as possible. eg "South African police cars are yellow" is more likely to be reliable than "All police cars are yellow".

2. Causality

Much of our reasoning about the world is not just a matter of generalisation (eg all dogs bark) but about relationships of cause and effect (eg smoking causes lung cancer; the effect of not studying is to fail). How can we make inferences about causal relationships without falling into the Fallacy of Post Hoc Reasoning? What sort of evidence is needed to support the claim that A causes B?

The question "what causes X to occur" can be split into two parts:
i) What factors must be present in order for X to occur?
ii) What factors invariably result in X occurring?

For instance, oxygen must be present in order for wood to burn, however, wood will not invariably burn in the presence of oxygen. If wood is made hot enough in the presence of oxygen then it will invariably burn.

Definition: Necessary and Sufficient Conditions
A necessary condition for the occurrence of a specified event is a circumstance in whose absence the event cannot occur.
A sufficient condition for the occurrence of a specified event is a circumstance in whose presence the event must occur.

The word "cause" sometimes means "necessary condition" and sometimes "sufficient condition". Often a cause is some circumstance (or set or circumstances) which is both a necessary and sufficient condition.

Consider any proposition in the form "if P then Q". Such a proposition could be phrased equivalently as "Q only if P" or "P is a sufficient condition for Q" or "Q is a necessary condition for P". If P is both a necessary and sufficient condition for Q, we often write "P if and only if Q" or simply "P iff Q".

Exercises
1. Is oxygen a necessary or a sufficient condition for fire, or both, or neither?
2. Is drowning a necessary or a sufficient condition for a person dying, or both, or neither?
3. Is validity a necessary or sufficient condition for soundness, or both, or neither?
4. Is soundness a necessary or sufficient condition for validity, or both, or neither?

Mills' Methods

John Stuart Mills formulated five methods of establishing whether events were causally related. Some of these principles were known before his 19th century formulation, but his name has become attached to them anyway.

1. The Method of Agreement

Suppose we have studied the circumstances under which an event E occurs and find the following--

Events prior to E on six separate occasions
A B C D
A B F G H
A C I J
A C I J K
A B F G H
A B C D F

It would seem likely that A is the cause of E, since A is the only point of agreement between the six cases.

> Definition: Method of Agreement
> If one (and only one) relevant circumstance is common to all cases in which an event occurs, then this circumstance is the cause of the effect (or at least is related to the cause).

> Exercises
> 1. Thembi has had a stomach ache for the last five days and thinks that it has something to do with the breakfast she eats.
>   Monday - Banana, milk, sugar, egg, coffee
>   Tuesday - Tea, milk, toast, banana, butter
>   Wednesday - Fish, egg, toast, butter, banana
>   Thursday - Meusli, milk, banana, orange juice
>   Friday - Banana, orange juice, toast, butter, milk
> 2. Lightning always occurs before thunder. Is it accurate to infer by the Method of Agreement that lightning causes thunder?

Clearly, the Method of Agreement is not 100% reliable. How can you be sure that you have taken into account all relevant circumstances? Consider the following example - Themba has been drunk for the last 3 nights. On Monday he drank Scotch and soda; on Tuesday he drank Whiskey and soda; and last night he drank Vodka and soda. What causes his drunkenness?

2. The Method of Difference

Again, suppose we have studied the circumstances under which an event E occurs and find the following -

Relevant Circumstances	Effect
A B C D F	E occurs
B C D F	E does not occur

It seems that when A is present (along with other factors) E occurs, but when A is absent and everything else is the same, E does not occur.

> Definition: Method of Difference
> If two situations are alike in all relevant respects except one, and the effect occurs in one instance but not the other, then the one difference is the cause of the effect (or at least is related to the cause).

3. Joint Method of Agreement and Difference

This is simply a combination of the two previous methods.

Relevant Circumstances	Effect
A B C	E occurs
A C D F	E occurs
C D F	E doesn't occur
B C	E doesn't occur

Once more, A looks like the cause of E.

4. Method of Concomitant Variation

Often causation is more complex than events either happening or not happening: we may need to take into account the degree or extent to which something occurs. For instance, it is not drinking vodka which makes one drunk, it is drinking too much vodka. The degree of drunkenness varies in proportion to the amount consumed.

Relevant Circumstances	Effect
A B C	E
A↑ B C	E↑
A↓ B C	E↓

> Definition: Method of Concomitant Variation
> When one circumstance varies in a regular manner whenever some other circumstance varies, then the two must be causally connected.

Note that this principle does not establish which event is the cause and which is the effect, or whether both are caused by some third factor.

Note also that the direction of variation does not matter (both increasing, both decreasing, or one increasing while the other decreases), as long as it is a consistent variation.

5. Method of Residues

We also want to account for events which have a number of joint causes. Suppose circumstances A,B and C consistently give rise to some event E and suppose that we know A is a cause of part of E. Then either B or C (or both) probably cause the rest of E.

> Definition: Method of Residues
> Where part of an effect remains unexplained by known causal circumstances, then additional circumstances must be sought to account for the unexplained portion of the effect.

eg A person drives a truck full of sand onto a weighing platform and the scale registers 5.4 tonnes. We know that the truck on its own causes the scale to register 1.2 tonnes and hence may infer that the driver and the sand must account for the other 4.2 tonnes.

3. Hypothesis formation and refutation

The Principle of Generalisation and Mills' Methods for studying causation all depend on observations. But many of our explanations of why things happen as they do require more than observation. An essential principle in both science and everyday life is the development of hypotheses to explain observed phenomena.

To hypothesise is to form an explanation for something. A hypothesis (or theory) proposes some unobserved phenomena to account for some observed phenomena.

eg We observe many objects falling to the ground and could form a generalisation such as "all objects fall to the ground unless some force intervenes". However, this generalisation does not explain why this should be the case. We need to go beyond the observations and propose that objects fall because of a force of attraction between objects (which we call gravity) whose strength is proportional to the mass of the two objects and the square of the distance between them. No-one has observed gravity, but the hypothesis of gravity is used to explain the observations of falling objects.

Abduction

Hypotheses are formed using the Principle of Abduction.

Definition: Principle of Abduction	
D is true	(D for "data")
H would explain D	(H for "hypothesis")
∴ H is true	

This is an important and common principle, yet we cannot guarantee its conclusions. It is possible for the premises to be true but the conclusion false and hence this form of argument is not logically valid. In fact abduction is a form of the Fallacy of Affirming the Consequent (D, if H then D ∴ H).

Consider the following example - Themba's girlfriend Alice has dumped him. Now, if Alice had discovered that Themba was married that would explain why she dumped him. But we cannot immediately infer that it must be the case that Alice discovered that Themba is married: she could equally well have left him for any number of other reasons.

Justification of a Hypothesis

Abduction may not be 100% reliable, but nevertheless some hypotheses are more reliable than others. How does one judge between hypotheses and decide which is most reasonable?

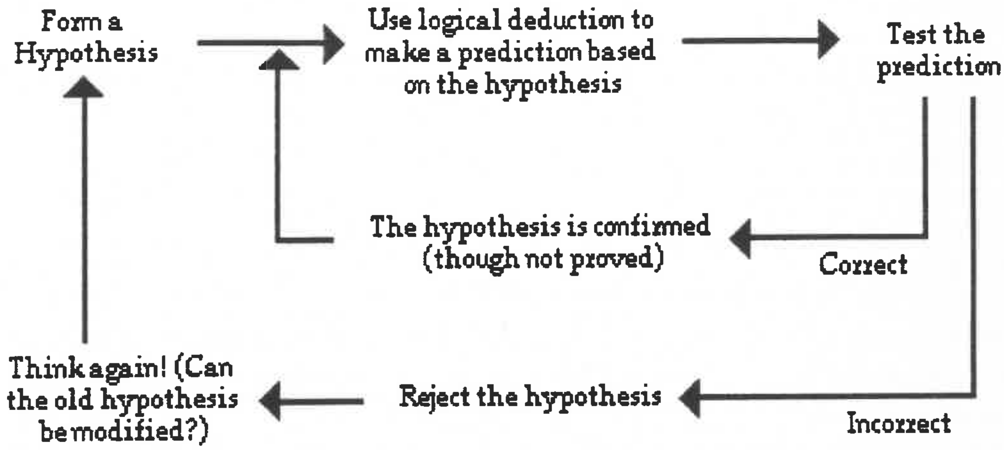
1. There should be many observations which are consistent with the hypothesis.
2. There should be a wide variety of observations which are consistent with the hypothesis.
3. There should be no observations which contradict the hypothesis[1].
4. The hypothesis should explain rather than just imply the data.
   eg The hypothesis "unsupported objects fall to the ground" certainly implies that if you let go of your pen it will fall, but it does nothing to explain why.
5. The hypothesis should be internally consistent
   eg "I robbed Mrs Smith because she owed me money and furthermore I was in Durban at the time so how could I have done it?" is not a very good explanation.
6. The hypothesis should be the simplest of all available hypotheses; the one which requires the least number of assumptions (often called Ockham's Razor or the Law of Parsimony).
7. Falsifiability - it must be possible to test the hypothesis; it must be clear what sort of observation would prove the hypothesis to be false.
8. The hypothesis should be able to generate predictions.

---

[1] At this stage we use the term "contradict" to mean that it is impossible for both the observation and the hypothesis to both be true. Later we will need to be more precise in our use of this word.

The Hypothetico-Deductive Method

Points 7 and 8 above are linked together in a process which forms the basis of scientific discovery.



Notice that this process is infinite: experiments may confirm a hypothesis but can never make it 100% guaranteed. However, just <u>one</u> solid fact which is inconsistent with the hypothesis is enough to falsify (disprove) it.

This process is also relevant to the task of debugging computer programs. At university, most programs are small enough that they can be exhaustively tested, but in the real world software is typically so complex that it is impossible in practice to ever guarantee that a system does what it was designed to do. There are mathematical (ie deductive) techniques to prove that a program is correct, but these cannot be applied in practice to any program over a couple of pages in length. In practice, checking that a program works correctly is an inductive process.

One starts with the hypothesis that the program is working as it was designed to and from that hypothesis predicts how the program will react to particular input. The program is then run using that input and the program's behaviour is compared to the predicted behaviour. If the prediction was correct then our confidence in the program increases, but it does not guarantee that the program is without fault. In order to reach such certainty, <u>every</u> possible input must be tested, and if the actual behaviour ever differs from the designed behaviour then we know that there is still a bug to be fixed. The task of testing every possibility is generally unfeasible, and so we must make do with some limited number of tests. It would take a huge number of tests (possibly infinite!) to show that a program is 100% correct, but only one failed test to show that the program is faulty.

The choice of tests then becomes crucial. It is totally inadequate to only test a computer program with input which you know will work. In the task of checking and debugging a program the tester must be ruthlessly committed to finding errors. Tests must be designed to push the program to its limits. The purpose of the hypothetico-deductive method is to continually try to falsify the hypothesis by running tests which will uncover counter-examples to the hypothesis.

4. Analogy

Consider the following two stories -

<u>Story 1</u> - A large army plans to besiege a city on a hill. The army is so large that coordinating transportation of supplies and troop movements along the narrow roads is a major problem. If the whole army approaches the city together the logistic difficulties will probably cause their attack to be ineffective. But sending a smaller force would be too weak to win the ensuing battle. The General in charge decides to split the army into three parties which will approach the city

from three different directions simultaneously. This will decrease the logistics difficulties as well as allowing the army to attack the city at three separate points.

Story 2 - A doctor is puzzling over how to treat a patient with a brain tumour. X-ray therapy is certainly the best option, but if a large beam of x-rays is directed at the tumour, intervening brain cells will also be damaged. However, if a lower-intensity beam is used the tumour will not be destroyed.

Being a clever doctor who remembered her military history, she reasoned that three separate low intensity beams directed at the tumour from different directions would destroy the tumour yet not damage the brain tissue. This is an example of reasoning by analogy: the doctor drew a principle from one setting and applied it in a different setting.

Analogy is not always used as a method of argumentation: "My wife's smile is like a ray of sunshine" is meant to be descriptive, not a proof of some conclusion. But it is often the case that we argue by analogy, indeed most of our everyday inferences are based on analogy.

- eg1 When we say "I'm going to buy another pair of Nike takkies because the last pair I bought were excellent", we are reasoning by analogy. We are in effect saying that one pair of Nikes is similar to another and so if one is excellent, so will another.
- eg2 When we rent a car for the first time we may recall our previous experience of buying a car and renting a house. These previous experiences help us to know how to deal with the new experience.
- eg3 The use of precedents in court cases.
- eg4 When the kettle broke, the problem was a loose wire in the plug. Now that my computer isn't working, I wonder whether it also might be a problem with the plug.

Definition: Reasoning by Analogy

A is like B
A has some property or quality x
∴ B also has x

Reasoning by analogy is probably the weakest form of induction. It is certainly not logically valid, but we can again ask what makes a conclusion based on analogy reasonable.

Exercise

Which of the following arguments is most reasonable? Why?

1. The first glass of wine from a bottle is like the second glass
   I enjoyed the first glass
   ∴ I will enjoy the second glass

2. 1981 Nederburg Pinotage is like 1985 Delair Cabernet
   I enjoyed 1981 Nederburg Pinotage
   ∴ I will enjoy 1985 Delair Cabernet

What do we mean by "A is like B"?

i) There are a number of similarities between A and B which are relevant to the conclusion
ii) There are no differences between A and B which would be relevant to the conclusion.

eg Medical research often experiments on animals and then applies the results to humans. Is this logically reasonable? Suppose drug D has a certain effect on a guinea pig. Is a guinea pig similar enough to a human to make an argument by analogy reasonable?

Similarities - both mammals and hence both have warm blood, heart, lungs, kidneys, liver, nervous system

Differences
- size (but maybe this will only affect dosage)
- guinea pig is herbivorous (certainly relevant if the drug has to do with digestion)
- guinea pig can't speak
- plus many others

Judging the Reliability of Conclusions

1. The more cases which fit the analogy, the more reliable the conclusion. This in effect changes the basic Principle to -

   A_1, A_2, A_3... are all like B
   A_1, A_2, A_3... all have some property or quality x
   ∴ B also has x

2. The cases A_1, A_2, A_3... should be as dissimilar as possible.
   eg Situation 1 - Drug D has been tried on 10 guinea pigs
   Situation 2 - Drug D has been tried on 10 animals including guinea pigs, monkeys and dogs.
   Conclusions about the effect of the drug on humans would be more trustworthy if based on the second situation than on the first.

3. The strength of the analogy, taking account of both the similarities between A_1, A_2, A_3... and B, and the differences.

4. The strength of the conclusion relative to the premises. The more specific the conclusion, the less reliable it becomes.
   eg My friend and I have similar handwriting
   My friend fits 800 words to a page
   The conclusion "∴ I could fit up to 800 words to a page" is more reliable than "∴ I will fit exactly 800 words to a page"

5. Relevance of the analogy - are the areas of similarity somehow connected to the property being concluded? Any two objects have some similarities, but the similarities may have no relevance to the argument.

Summary and Conclusion

1. Inductive logic includes various methods of reasoning which allow us to infer conclusions based on observations to some degree of probability.

2. None of the forms of induction we have examined (generalisation, causation, hypotheses and analogy) can guarantee their conclusions since they are not logically valid. Nevertheless they constitute the bulk of human reasoning.

3. If not used correctly, induction can result in fallacies such as hasty generalisation, post hoc reasoning and affirming the consequent.

4. Each of these forms of reasoning can be implemented on computer (although not as easily as the deductive methods in the next section): programs which diagnose medical or mechanical faults often use abduction; of growing interest in computer science is the field of Case-Based Reasoning which is a form of reasoning by analogy. Thus computers can not only do maths and formal deduction with guaranteed accuracy, but can also reason about uncertainty.

Section 3 - Axiomatic Systems

An axiom is a way of stating a symbolic definition. Axioms normally express some self-evident, uncontroversial truth. It is common in maths and logic to propose a set of axioms as the foundation on which to build a deductive system. For instance set theory, arithmetic and geometry are all axiomatic systems.

We try to make the set of axioms as small and as simple as possible and then derive everything else we need from the axioms. Anything which can be derived from the axioms is called a theorem. Theorems cannot be called true or false; we can only claim that they are derivable from the axioms.

Axioms and theorems are technically meaningless string of symbols. An axiomatic system defines "legal" ways in which those symbols can be manipulated. Within the system, the symbols have no meaning; it is only from outside the system that we interpret the symbols in some meaningful way. (This may be compared to something like Morse Code: a meaningless sequence of dots and dashes becomes meaningful when someone defines an appropriate interpretation.)

Example (Suggested by Douglas Hofstadter in "Gödel, Escher, Bach")
Consider an axiomatic system which contains only the symbols p, q and -. The rules defining how these symbols can be manipulated are as follows -

1. If x represents any string of hyphens, then xp-qx- is an axiom.
   eg -p-q-- is an axiom
   ----p-q----- is an axiom
2. Suppose x, y and z all represent strings of hyphens.
   If xpyqz is a theorem then so is xpy-qz-.

(Note that by definition, any axiom is also a theorem.)

What sort of theorems can be derived in this system?

T_1:	-p-q--	from Rule 1
T_2:	-p--q---	from T_1 and Rule 2
T_3:	-p---q----	from T_2 and Rule 2
etc		

Exercises

1. Are the following strings theorems?
   i) ---p-q----
   ii) ---p----q-------
   iii) --q-
   iv) --p--p-q-----
   v) --p---q-
2. Is there some general decision procedure which will tell us easily whether some string is a theorem or not?

But what does a string like --p-q--- mean? Technically, nothing; these are just strings of meaningless symbols. However, maybe from outside the system we can give the symbols some meaning.

i) We could interpret - as representing "apple", p as representing "horse" and q as representing "happy". Then --p-q--- means "apple apple horse apple happy apple apple apple". But this interpretation isn't very satisfying since theorems make no more sense than non-theorems.

ii) What if - represents "1", -- represents "2", etc, p represents "plus" and q represents "equals"? Under this interpretation --p-q--- means "2 plus 1 equals 3". Under this interpretation, every theorem corresponds to some valid arithmetic expression.

iii) But don't think that ii) is the only meaningful interpretation of the system: there may be an number of others. For instance, p might mean "equals" and q might mean "taken from".

Computers and Logic

Computers work in much the same way - they manipulate patterns of high and low voltage which, to the computer, have absolutely no meaning. The voltages do not even "mean" 0 and 1! But we interpret a low voltage as 0 and a high voltage as 1; then we interpret strings of 0's and 1's as binary numbers; then we create something like the ASCII code to arbitrarily assign bit-patterns to A, B, +, $, a, b etc. Furthermore we program the computer to display certain pixel patterns on a VDU in ways which we then interpret to mean letters, numbers etc.

This may seem a trivial point, but only because we have become so used to this way of interpreting the computer's behaviour. A computer is purely a syntactic processor which manipulates symbols according to axiomatically defined procedures. Computers are strictly logical, and logic itself can be expressed as an axiomatic system (see the end of Section 5).

Exercise

Suppose x and y represent some combinations of the letters "M", "I" and "U" (eg "MIIU", or "MMMIMU", or "UI". Note that "" is a valid combination of zero letters.).

Axiom rule 1 -	The string of letters "MI" is a theorem
Axiom rule 2 -	If some string "xI" is a theorem then "xIU" is also a theorem (eg since MI is a theorem, so is MIU)
Axiom rule 3 -	If "Mx" is a theorem then "Mxx" is also a theorem
Axiom rule 4 -	If "xIIIy" is a theorem then "xUy" is also a theorem
Axiom rule 5 -	If "xUUy" is a theorem then "xy" is also a theorem

Either derive the following theorems or prove that they cannot be derived -

i) MIIII
ii) MIIMUU
iii) MUIUIU
iv) MIIIII

Section 4 - Sets

Definitions

By a "set" we mean a collection of any sort (eg the collection of all university students; a pair of shoes; a bouquet of flowers; a flock of sheep; the three primary colours ...). It is not necessary for the members/ elements of a set to have anything in common, though interesting sets are those whose elements are somehow related to each other. The terms class, aggregate and collection are synonymous with "set".

Set membership

If p is an element of a set A then we write $p \in A$, otherwise $p \notin A$.

Well defined

A membership condition (also called a predicate) is some rule which specifies which objects are members of a particular set. A set is said to be well defined if it has a membership condition.

eg PR = {x | x is prime and x < 24} is well defined
F = {all famous people} is not well defined

Enumeration

An alternate way to define a set is by enumerating (ie listing) all its members.

eg PR = {2,3,5,7,11,13,17,19,23}

Cardinality

$|A|$ denotes the number of distinguishable elements in a set A. A set is said to be "finite" if it's cardinality is finite, otherwise it is "infinite".

eg N = {0,1,2,3,4, ...} is the infinite set of natural numbers.

Sets of sets

Sets may contain other sets, eg A = {a, {a,b}, c, {b,c}, {d}, d}. Note that d is not the same as {d}. In the above example, $d \in A$, $d \in \{d\}$ and $\{d\} \in A$. $|A| = 6$.

It is also possible (though unusual) for a set to be a member of itself. For instance, consider the set $S = \{A \mid |A|>7\}$, that is the set of all sets with more than seven elements. Now $|S|$ is certainly larger than 7 (in fact it is probably infinite) and hence $S \in S$.

Subsets

$A \subseteq B$ (read as "A is a subset of B") iff for every $a \in A$ it is also true that $a \in B$. The subset relation is reflexive (ie $A \subseteq A$) and transitive (ie if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$). If $A \subseteq B$ then $|A| \leq |B|$.

Equivalence

$A \equiv B$ ("A is equivalent to B") iff $A \subseteq B$ and $B \subseteq A$. That is, two sets are equivalent iff they contain exactly the same elements. From this definition we can see that duplicates are not significant and are therefore omitted. You will see for instance that the following sets are equivalent - {a,a,b,a,a,c,a}, {a,b,c} and {c,c,c,b,b,a}.

(Because I'm lazy I often use = and ≡ interchangeably.)

Proper subset

$A \subset B$ iff $A \subseteq B$ and $A \neq B$. This relation is transitive but not reflexive.

Universal set

The universal set (normally called U or E) includes every object under discussion.

Empty set

$\emptyset$ (sometimes written {}) is the empty (or null) set which contains no elements at all. $|\emptyset|=0$.

If you look closely at the previous definitions, you'll see that for any set A, $\emptyset \subseteq A$. Don't confuse this with $\emptyset \in A$, which is not always true.

Power set

For any set A, we can derive the power set $p(A)=2^A=\{x \mid x \subseteq A\}$. Note that according to this defintion, the empty set and the set A itself are both members of p(A). If A is finite and $|A|=n$ then $|2^A| = 2^n$.

Intersection

$A \cap B=\{x \mid x \in A \text{ and } x \in B\}$

Notice that
$A \cap B = B \cap A$ (Commutativity)
$A \cap A = A$
$A \cap \emptyset = \emptyset$
$A \cap U = A$
$(A \cap B) \cap C = A \cap (B \cap C)$ (Associativity)
$A \subseteq B$ iff $A \cap B = A$

Disjoint

A and B are disjoint iff $A \cap B=\emptyset$. A collection of sets are said to be disjoint iff every pair of sets in the collection are disjoint.

Union

$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$

Notice that
$A \cup B = B \cup A$ (Commutativity)
$A \cup \emptyset = A$
$A \cup A = A$
$(A \cup B) \cup C = A \cup (B \cup C)$ (Associativity)

Complement

$\overline{A}=\{x \mid x \notin A\}$ (may also be written as ~A)

Occasionally it is useful to define a relative complement between two sets. A\B is read "A without B" or "A except for B" and is defined as $A \setminus B = \{x \mid x \in A \text{ and } x \notin B\}$. With this definition it is clear that $\overline{A}=U \setminus A$.

Notice that if $A \subseteq B$ then $A \setminus B=\emptyset$.

Exercises

1. Enumerate all the elements of the following sets (in all cases the universe of discourse is the set of integers).
   i) $\{n \mid 4<n<9\}$
   ii) $\{n^2 \mid 4<n<9\}$
   iii) $\{n^2 \mid 4 \le n \le 9\}$
   iv) $\{p \mid p \text{ is prime}\}$
   v) $\{x^3+1 \mid -2 \le x \le 2\}$

<u>Exercises (continued)</u>

2. Let the universe U be all the non-negative integers less than 10 and let $A = \{1,2,3,6\}$, $B = \{n \mid n \text{ is odd}\}$ and $C = \{n \mid n \text{ is not divisible by } 3\}$. Represent this situation with a Venn Diagram and list the elements in the following sets -
   - i) $\sim C$
   - ii) $A \cup B$
   - iii) $\sim A \cup B$
   - iv) $\sim(A \cup B)$
   - v) $A \cup A$
   - vi) $A \cap A$
   - vii) $A \cap (B \cap C)$
   - viii) $(A \cap B) \cap C$
   - ix) $A \cup (B \cap C)$
   - x) $(A \cup B) \cap C$
3. Let C be the set of airplanes that made a landing on the moon between 1850 and 1900. Let D be the set of refrigerators sold to South African citizens who live within 100km of the North Pole. Are these sets equivalent? Disjoint?
4. If $A = \{R,G,B\}$, list 2^A. What is $|2^A|$?
5. Imagine a simple colour computer monitor where each pixel contains three sub-pixels (a red, a green and a blue) which each may be either on or off at any particular time. How many different colours could be displayed on such a screen?
6. Let A be the set of positive integers and B the set of positive real numbers. What is the cardinality of these sets. Does A contain the same number of elements as B?

Various set identities

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributive Laws)
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

$\sim(A \cup B) = \sim A \cap \sim B$ (de Morgan's Laws)
$\sim(A \cap B) = \sim A \cup \sim B$

$\sim\varnothing = U$
$\sim U = \varnothing$
$A \cup \sim A = U$
$A \cap \sim A = \varnothing$

$|A \cup B| = |A| + |B| - |A \cap B|$

Ordered pairs

An ordered pair consists of two objects in a definite fixed order and is denoted by <x,y>. Note that order is important: <x,y>≠<y,x>. If P=<x,y> and Q=<u,v> then P=Q iff x=u and y=v. (You should already be familiar with this concept since it is used in Cartesian Geometry, except that round brackets are used instead of angle brackets.)

We can talk about ordered triples, quadruples etc in the same way. The general *n-tuple* is written $\langle x_1, x_2, \ldots x_n \rangle$.

Cartesian product

If A and B are sets then $A\text{x}B = \{\langle x,y \rangle \mid x \in A \text{ and } y \in B\}$. AxB is normally read "A cross B" and hence the Cartesian Product is often called the <u>cross product</u>.

Notice that
$A\text{x}\varnothing = \varnothing\text{x}A = \varnothing$
$A\text{x}B \neq B\text{x}A$ unless $A \equiv B$ (not commutative)
$A\text{x}(B\text{x}C) \neq (A\text{x}B)\text{x}C$ (not associative)
$|A|.|B| = |A\text{x}B|$
$A\text{x}(B \cup C) = (A\text{x}B) \cup (A\text{x}C)$
$A\text{x}(B \cap C) = (A\text{x}B) \cap (A\text{x}C)$

Exercises

1. Draw Venn Diagrams to illustrate each of the set identities listed above.
2. A Venn Diagram with three circles has 8 (why does this number keep coming up?) separate regions. Find a set expression to define each region.
3. Draw Venn Diagrams to represent the five possible arrangements of two sets. (Two overlapping circles can be used to represent all five in one diagram, but the task here is to represent each possibility separately.)
4. If A = {R,G,B} and C = {-1,1}, list the elements of AxC. What is |AxC|?
5. In a class of 20 students, each student is marked out of 100. If C is the set of students and M is the set of possible marks, what does CxM represent? What is |CxM|?
6. (Difficult) All educated people now realise that pprills, squearths and glops have all proved to be simply forms of nuph. It is also well established that squearths are both glops *and* nuphs. However, there is a complication. Recent work has established that there are glops which are neither squearths, gdynxs nor pprills. Further there are squearths which are neither gdynxs nor yet pprills.
   Admittedly some pprills *are* glops as are all squearths and even some gdynxs as well. Now that we know more about gdynxs, that some are squearths, some glops and some, unfortunately, both pprills *and* squearths, there are certain urgent questions that can be answered definitively.
   i) Can the universe contain such an unfortunate creature that, to be truthful, it must admit that it is a pprill, a nuph, a squearth, a glop and also a gdynx?
   ii) Consider those gdynxs which are not nuphs, can they possibly be glops?
   iii) If a pprill is a squearth, is it also a glop and can it possibly be a gdynx?

What is the status of set theory?

These basic concepts seem to be the most simple and uncontroversial mathematical concepts possible. However, three important points need to be noted:

i) Far from being simple, set theory forms the theoretical basis for virtually all maths, logic and computing.

ii) Set theory can be expressed as an axiomatic system. There is much controversy over the best collection of axioms, but the nine Zermelo-Fraenkel Axioms are the most commonly used.

iii) Although the individual concepts relating to sets seem indisputably true, when combined they give rise to some surprisingly troublesome results, such as the following -

Russell's Paradox

In 1901, Bertrand Russell presented the following paradox to show that some difficulties exist with these "simple" set concepts.

Assume U is the universal set which contains absolutely everything and consider the following sets:

$$V = \{x \in U \mid x \in x\}$$
$$W = \sim V = \{x \in U \mid x \notin x\}$$

The crucial question now is whether or not $W \in W$.

Case 1 - suppose $W \in W$. Then by definition, $W \in \{x \in U \mid x \notin x\}$ and hence $W \notin W$ (which is a contradiction).

Case 2 - suppose $W \notin W$. Then $W \in V$, which means that $W \in \{x \in U \mid x \in x\}$, and hence $W \in W$ (another contradiction).

In summary, if $W \in W$ then $W \notin W$ and if $W \notin W$ then $W \in W$! Since both possibilities lead to contradictions it must be that our original assumption was wrong. Thus, the conclusion from Russell's Paradox is that there cannot be a truly universal set.

Berry's Paradox

Consider the set Q = {Integers not nameable in fewer than nineteen syllables}. eg 1,362,795$\in$Q since it takes 20 syllables to pronounce. Now consider the smallest number in Q. Maybe this number is 111,777, but whatever it is just call it α.

By definition $\alpha \in$Q, but it can be uniquely identified in 18 syllables ("the least integer not nameable in fewer than nineteen syllables") and hence is not a member of Q!

This is a bit more serious than Russell's Paradox because it points out a problem with the way we choose objects out of a set and give them names. There are deeper difficulties too which show that there is something not quite right about set theory, but nobody is quite sure where the problem lies. For interesting reading on this, see "Infinity and the Mind" by R. Bucker (pub. Birkhäusser, 1982).

Section 5 - Propositional Logic

There are various ways to formalise logic: syllogistic logic, propositional logic, predicate logic, modal logic, fuzzy logic .... These vary in their complexity and their range of usefulness. In Propositional Logic, every proposition is seen as an atomistic claim with no internal structure. For instance the propositions "some people are tall" and "there are no unicorns" will both be treated the same. The simplicity of this assumption makes it easy to learn, but restricts its usefulness.

Any proposition may be represented by a symbol called a propositional variable. eg C may represent the proposition "Cheese is always yellow". A simple statement is one which contains only one proposition, whereas a compound statement contains a number of propositional variables combined with the following connectives -

Technical Name	Common Name	Symbol
Negation	not	~ (or ¬)
Conjunction	and	& (or ∧)
Disjunction	or	v
Material Implication	implies	⇒ (or ⊃)
Bi-conditional	iff	⇔

Set Theory and Propositional Logic

One of the characteristics which makes propositional logic easy to understand is that it closely parallels set theory.

English	Set Theory	Propositional Logic
1. "It is raining today"	Membership. Let R be the set of rainy days. The statement claims that today$\in$R .	A single proposition represented by a variable, say R. The statement claims that R is true.
2. "It is not raining today"	Complementation. today$\notin$R	Negation. The statement claims that R is false. In other words ~R is true.
3. "It is either raining today or I am blind"	Union. If B is the set of blind people, then the statement expresses R$\cup$B.	Disjunction. The statement is a combination of two propositions connected by the word "or". This would be written as RvB where "v" comes from the Latin word "vel".
4. "It is rainy and cold today"	Intersection. If C is the set of cold days then the statement claims that today$\in$R$\cap$C.	Conjunction. The statement is a combination of two propositions connected by the word "and". This would be written as R&C.
5. "If it is raining today then the grass will grow"	Subset. If G is the set of situations in which grass grows, then the statement says R$\subseteq$G.	Material Implication. The antecedent R implies the consequent G - written as R⇒G.

6. "All people with <50% will fail and the only way to fail is to get <50%"	Equivalence. If L is the set of people with low marks and F is the set of people who fail, then this statement claims that L≡F.	Bi-conditional. L is true iff F is true. This is written as L⇔F.

Well-formed Formulae

A sentence in propositional logic is any string of brackets, variables and connectives.

eg S&(RvT)
P
A&v)F

Notice that not all sentences are sensible, so we need some way of defining the set of sensible sentences -

> Definition: Well-Formed Formulae
> A well-formed formula[1] (wff) is a sentence which meets the following restrictions -
> i) Any propositional variable on its own is a wff; and
> ii) If A and B are wffs then so are ~A, (A&B), (AvB), (A⇒B) and (A⇔B).

By convention, it is permissible to omit the outermost pair of brackets. eg one could write P⇒(Q&R) rather than (P⇒(Q&R)). However, all other brackets are important to avoid ambiguity.

> Exercise: Ambiguity
> Think about the two ways in which the sentence P⇒Qv~A could be interpreted. Which one most suits the statement "If your car won't start, either catch a lift with a friend or don't go to university at all".

From this simple recursive definition, an infinite number of wffs may be generated[2]. Notice that the definition requires that the operators &,v,⇒ and ⇔ always have exactly two operands. Thus, a sentence like (P&Q&R) is not a wff: you must write either (P&(Q&R)) or ((P&Q)&R).

1 Don't be confused by the different spellings "formula" and "formulae" - the latter is simply the plural of the former.

2 This parallels the way in which "grammars" are used to define computer languages. For instance the Pascal grammar is a set of recursive rules from which an infinite variety of Pascal programs may be derived.

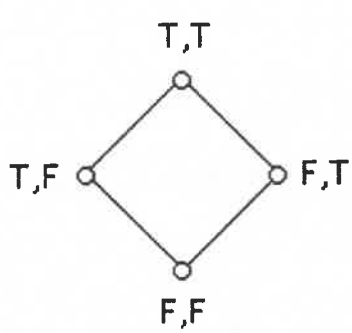
Exercises: wffs

1. Which of the following sentences are wffs?
   - i) ~~~~P
   - ii) P~Q
   - iii) (P,Q,R)
   - iv) (~(A)&B)
   - v) ~((A)&B)
   - vi) (AvBvC)
   - vii) {x|x∈A}
   - viii) (A⇒B)⇔(B⇒A
   - ix) (P⇒(Q⇒(~R⇒((S&T)⇔P))))
2. Represent the following as propositional formulae -
   - i) Mandy has blue eyes
   - ii) Bob is a football player and goes to the gym
   - iii) Mandy wants to be fit but never goes to the gym
   - iv) If Bob met Mandy they would fall in love or I'll eat my hat
   - v) Bob met Mandy and they didn't fall in love
   - vi) Blued eyed boys are attractive
   - vii) If I meet a blue-eyed girl, I am attracted to her and if I'm attracted to someone, they always turn out to be a blue-eyed girl.

Possible Models Diagrams

The important question which propositional logic seeks to answer is "under what circumstances is a wff true and under what circumstances is it false?" eg When is the wff (P&Q)⇒(QvP) true? Suppose P represents "1+1=3" (which turns out to be false) and Q represents "The Lecturer's name is Matthew" (which is true). Under those circumstances (P&Q) is false, but (QvP) is true. What about the truth of the wff as a whole? Is (P&Q)⇒(QvP) true or false? Furthermore, in different circumstances, when P and Q are interpreted differently, will (P&Q)⇒(QvP) be true or false?

On the surface it would seem to be difficult to say much about (P&Q)⇒(QvP) in general since there are an infinite number of ways to interpret (P&Q)⇒(QvP) since P and Q could represent any two propositions. However, although there are an infinite number of interpretations of P and Q, they may be grouped into four categories. Regardless of how you interpret P, the proposition will either turn out to be true or it will turn out to be false; and likewise for Q. Hence there are four possible states of the world: the state in which both P and Q are true; the state where P is true but Q is false; the state where P is false but Q true; and the state where both P and Q are false. We could show these situations in a simple diagram -



Each of these possible situations is called a model of P and Q and hence the picture is called a Possible Models Diagram (PMD). We show that some expression is true for a particular model by filling in the corresponding corner of the PMD. Thus the propositions P and Q can be represented by the PMDs -

Combining Rules

Not only are Possible Models Diagrams an easy way of visualising simple statements, they can also be combined in various ways to represent compound statements.

1. Negation

If P is true then ~P must be false; and conversely if P is false then ~P must be true. In other words any model which satisfies P does not satisfy ~P, and no model which does not satisfy P does satisfy ~P. Hence to form a PMD for ~P we take the PMD for P and reverse each corner: if the corner is filled in, rub it out; if it is blank, then fill it in.

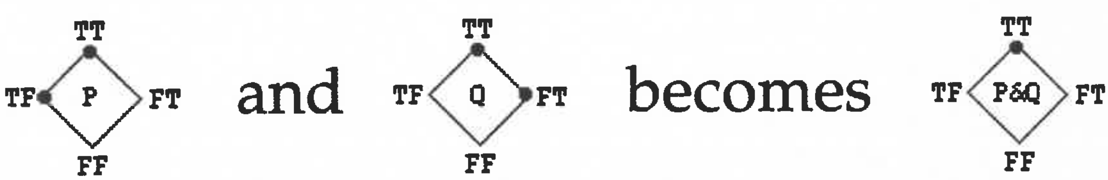
(Note the similarity with set complementation.)

Exercise: Negation

Suppose some wff α has the PMD . Draw the PMD for the wff ~α.

2. Conjunction

P&Q is a wff claiming that both P is true and also Q is true. So a PMD for P&Q will have to combine the diagram for P with the diagram for Q by matching-the-dots from two diagrams - if both original diagrams have a dot in a particular corner then the diagram of their conjunction will also have a dot in that corner.



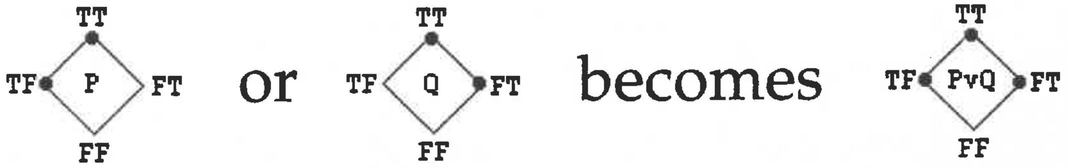
(Note the similarity with the set intersection operation.)

Exercise: Conjunction

Construct a PMD for the conjunction  & .

3. Disjunction

The wff PvQ is true whenever either P is true or Q is true or both are true. To form a PMD from the disjunction of two other PMDs, simply overlay the two original diagrams.



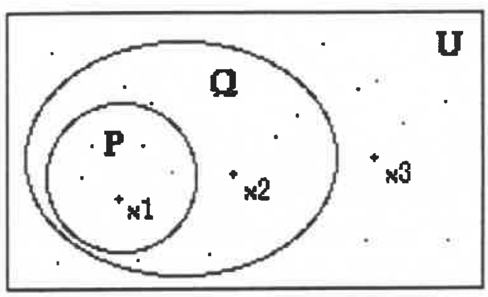
(Note the similarity between overlaying and the set union operation[1].)

In English we sometimes use "or" inclusively (eg "you can ride the roller-coaster if you're over 15 or accompanied by an adult") but other times exclusively (eg "either you eat your vegies or you won't get any pudding"). In logic **disjunction is always inclusive**.

Exercise: Disjunction
Construct a PMD for the disjunction  v .

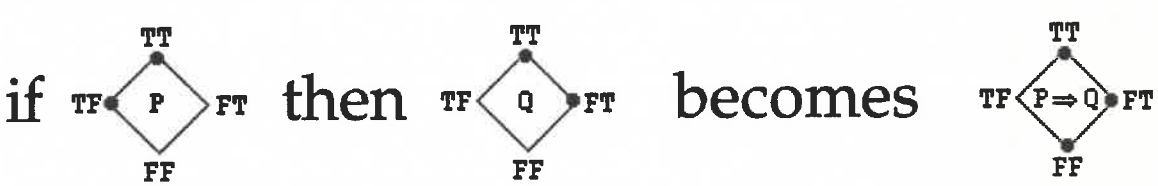
4. Material Implication

Remembering that the conditional is an attempt to capture the same concept as subset. So in order to form a PMD for P⇒Q, we first consider the analogous situation in set theory -



If P⊆Q, could there be a situation where x∈P and x∈Q?
Could there be a situation where x∉P and x∈Q?
Could there be a situation where x∉P and x∉Q?
Could there be a situation where x∈P and x∉Q?

x∈P and x∉Q is the only possible situation which disproves that P⊆Q. By analogy, in propositional logic, P and ~Q is the only model for which P⇒Q is false. Thus -



Two things to note about this diagram -

i) It is easy to remember that ⇒ has the diagram  because the three dots look like an arrow pointing to the right.
ii) When combining two diagrams as above, the visual procedure is to reverse the first diagram and overlay it on the second.

---

1 Take care not to confuse the terms "disjunct" and "disjoint". A disjunction is a logical concept which bears no relation to the concept of disjoint sets.

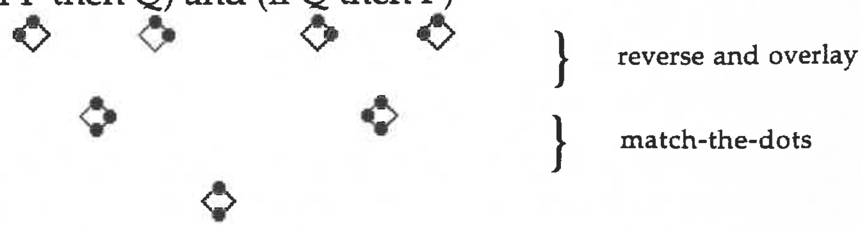
<u>Exercise: Implication</u>

Construct a PMD for the conditional .

5. Bi-conditional

The bi-conditional P⇔Q is defined as (P⇒Q)&(Q⇒P) and thus can be diagrammed by appropriate combinations of the previous operations -

(if P then Q) and (if Q then P)



The final outcome is intuitively sensible, since it shows those models in which the truth of P is exactly the same as the truth of Q.

<u>Exercise: Bi-conditional</u>

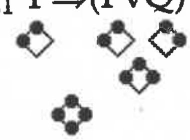
Construct a PMD for (P⇒Q)⇔(PvQ).

This is a bit time consuming because the definition of ⇔ makes the wff very long and there will be a lot of intermediate PMDs before you get to the final one. Can you think of a shortcut which would allow you to construct a diagram for ⇔ without needing to deal with such a long expression? That is, given two expressions (let's call them α and β) what visual procedure would you use to form a diagram for α⇔β?

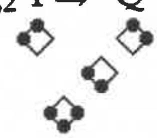
PMDs for longer wffs

Any wff which uses only two propositional variables can be represented by a PMD, and this diagram can be formed using the simple combining rules (reverse, overlay and match-the-dots) described above.

eg_1 P⇒(PvQ)



eg_2 P⇒~Q



eg_3 (PvQ)&~(PvQ)

<u>Exercise: PMDs</u>

1. Draw a PMD for the wff mentioned at the beginning of this section - (P&Q)⇒(QvP)
2. For each of the following wffs, try to find an interpretation which you consider makes the wff false and another which makes it true -
    - i) P&Q
    - ii) P⇒Q
    - iii) (P&~Q)⇒R
    - iv) (P&~Q)⇒P
    - v) P&~P

Truth categories for wffs

<u>Definition</u>
Every wff falls into one of three categories -

i) <u>Tautology</u> - a wff which is true regardless of how you interpret it. In propositional logic this means any wff whose possible models diagram has <u>every</u> corner filled in. (eg_1 above)

ii) <u>Contingent</u> - a wff which is sometimes true and sometimes false. In propositional logic, a wff whose possible models diagram has <u>some</u> corners filled in but not all. (eg_2)

iii) <u>Inconsistent</u> - a wff which is always false, regardless of what interpretation is given to it. In propositional logic, a wff whose possible models diagram has no corner filled in. (eg_3) (Sometimes also called a contradiction.)

In English, to say that something is a tautology is normally a criticism. If halfway through an argument someone claims that "all bachelors are male" we may say "of course, but that's just a tautology". ie It is just part of the definition of "bachelor" to be male; it is true, but trivial; to talk about a male bachelor is just to repeat oneself; the word "male" adds no information since bachelors are necessarily male. Tautologies are the most useful category in logic because they are <u>necessarily</u> true rather than <u>empirically</u> true. For instance it is necessarily true that Pv~P, since there is no conceivable interpretation of P for which Pv~P would be false.

<u>Exercises: Tautologies</u>
Draw PMDs to show that each of the following are tautologous -

i) ~(P&~P) The Law of Non-Contradiction
ii) Pv~P Law of Excluded Middle
iii) P⇒(QvP)
iv) (P&Q)⇒P
v) (P&Q)⇒Q

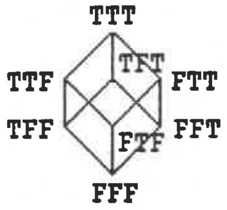
The term "contingent" indicates that the truth of the wff depends on the interpretation. Since the wff is true in some models yet false in others, if you want to know whether it is true in a particular situation you must hire a scientist to undertake an empirical study.

<u>Exercises: wffs</u>

1. Draw PMDs for each of the following wffs and classify them as tautologous, contingent or inconsistent -
   i) ~(P⇒(P&Q))
   ii) (PvQ)&~P
   iii) (PvQ)⇒(P&Q)
   iv) (P&~Q)⇒Q
   v) (Pv~P)⇒(P&~P)
   vi) ~(~Pv(PvQ))
2. For the PMD in part iii) above, write down 4 interpretations of P and Q, one for each of the four possible models.

Handling more than two propositional variables

If there are three propositional variables in a wff (say P,Q and R), there will be eight possible models. These are best pictured as the corners of a cube, but this can be shown diagramatically as the following graph (in which each triple shows the truth values of P, Q and R respectively) -



The positioning of the corners in this diagram is important. Notice that there are four rows - in the top row all variables are true, in the second row two are true and one false, in the third row two are false and one is true and in the bottom row all are false. The combining rules (reverse, overlay and match-the-dots) can all be used with this diagram as before.

Using truth tables (see below), a wff with three propositional variables would require eight rows. In general, if a wff contains n different propositional variables then there will be 2^n possible models. Hence, the wff's PMD would have 2^n nodes and its truth table would have 2^n rows. It is rare to use either PMDs or truth tables for wffs with more than three variables.

<u>Exercises: Three variables</u>
Draw PMDs for each of the following wffs and classify them as tautologous, contingent or inconsistent -

i) (P&Q)⇒R
ii) (P&(QvR))⇒((P&Q)v(P&R))
iii) ((P&Q)⇒R)v~P

Equivalence of wffs

<u>Definition: Equivalence</u>
Two wffs are equivalent iff they are true under exactly the same circumstances. ie A≡B whenever the PMDs for A and B are identical.

For instance (P⇔Q) ≡ ((P⇒Q)&(Q⇒P)) - in fact we have used this as the definition of the bi-conditional.

If A≡B then A⇔B will be a tautology.

Of course one could list an infinite number of equivalences, but some of the more common are -

de Morgan's Laws	~(P&Q) ≡ ~Pv~Q
	~(PvQ) ≡ ~P&~Q
Commutation	PvQ ≡ QvP
	P&Q ≡ Q&P
Association	Pv(QvR) ≡ (PvQ)vR
	P&(Q&R) ≡ (P&Q) &R
Distribution	P&(QvR) ≡ (P&Q)v(P&R)
	Pv(Q&R) ≡ (PvQ)&(PvR)
Double Negation	P ≡ ~~P
Transposition	P⇒Q ≡ ~Q⇒~P
Exportation	(P&Q)⇒R ≡ P⇒(Q⇒R)
Material Equivalence	P⇔Q ≡ (P&Q)v(~P&~Q)
Material Implication	P⇒Q ≡ ~PvQ

Exercises: Equivalence

1. Convince yourself that each of the pairs of wffs listed above are indeed equivalent.
2. We have said that the "v" connective represents inclusive-or. Find a wff (using some combination of the five connectives) which expresses the concept of exclusive-or.
3. i) Show how the connective ⇒ can always be replaced by some combination of ~ and v.
   ii) Similarly, show how the connective ⇒ can always be replaced by some combination of ~ and &.
   iii) Would it be possible to create a form of propositional logic which only uses the three connectives ~, & and v instead of all five which we use? If so, would it be useful?

Paradoxes of Material Implication

Special mention needs to be made about the last in the list of equivalences above. P⇒Q is equivalent to ~PvQ because of the way we derived the concept of implication from the subset concept, and it is this equivalence which justifies our use of "reverse and overlay" as the visual procedure for material implication.

There are some counter-intuitive aspects to this definition however. For instance, if you draw a PMD for the wff P⇒(Q⇒P), you will see that it is a tautology. This is a strange wff since it claims that if P is true then anything implies P - eg "The earth is round, therefore if the moon is make of cheese then the earth is round"!

Or consider the wff ~P⇒(P⇒Q) which claims that if P is false, then if P were true anything else would also be true - eg "Since computers don't have emotions, if computers did have emotions then Durban would be in Namibia". But this wff too is a tautology.

These two wffs are called paradoxes - there is nothing actually illogical or contradictory about them, it's just that they don't match our intuitions. These paradoxes occur because our definition of material implication does not exactly fit the way we often use the words "if ... then ...". In particular, the confusion arises because the conditional P⇒Q is defined to be true in the models where P is false.

Substitution

In logic we study the form of propositions and arguments without being concerned by what they. represent. Hence, we expect that the actual choice of variables names is immaterial. eg we expect P⇒Q to act the same as A⇒B. We don't care what symbol is used and we don't care what proposition the symbol represents.

Definition: Substitution

i) Let E be a wff containing the propositional variables $p_1, p_2, \ldots p_n$
ii) Let E* be obtained from E by substituting propositional variables $q_1, q_2, \ldots q_n$ simultaneously throughout E for $p_1, p_2, \ldots p_n$ respectively
iii) Then E* is called a substitution instance of E and E≡E*

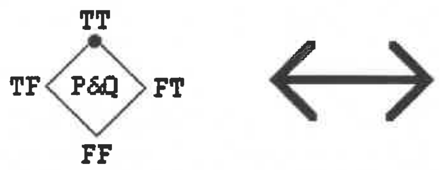
eg Suppose we have the wff (P&Q)⇒R
where P represents "the sun is shining
Q represents "there is no wind"
R represents "today is Friday"
Now substitute A and B for Q and R
where A represents "exams are in 9 weeks"
B represents "50 students are worried"
Then (P&Q)⇒R ≡ (P&A)⇒B (ie their PMDs will look exactly the same).

We shall have need later to extend this definition of substitution to allow a variable to be replaced not only by another variable, but by a whole wff. But this simplified form of substitution is all we require at present.

Truth Tables

Although PMDs are a useful visual technique for learning about logic, the <u>standard</u> way of describing wffs is the truth table. Truth tables show in tabular form the same information as a PMD and it is easy to convert from one to the other. Whereas the PMD represents the four possible models as corners of a square, the truth table represents them as separate lines in a table. Filling in a corner is equivalent to writing "T" in the appropriate row of the table.

eg Compare the PMD and the truth table for the wff P&Q -



P Q	P & Q
T T	T
T F	F
F T	F
F F	F

Truth tables are often written using 1's and 0's rather than T's and F's. This is especially true when truth tables are used in computing because it emphasises that boolean values can be used for both logic and arithmetic. In computer architecture courses you see electronic gates which act both as logical connectives and as mathematical operators. The electronic components of a computer work by differentiating voltage levels and this differentiation can readily be interpreted as either a boolean truth value or a binary digit. Not only are T's and F's replaced by 1's and 0's, but the truth table is often shown more like a multiplication table, to emphasise the dual roles as both numerical digits and logical values. eg -

&	0 1
0	0 0
1	0 1

(See the later section on multi-valued logics which shows how truth values can be manipulated mathematically.)

<u>Exercises: Truth Tables</u>

1. Draw truth tables for each of the five logical connectives.
2. Select a variety of wffs from previous exercises and convert the PMDs you have drawn into truth tables.
3. Select a variety of wffs from previous exercises and construct truth tables directly (ie without drawing PMDs first).

<u>Exercise - Shortcuts</u>

Drawing a PMD for a long wff requires a lot of intermediate diagrams, but now that you are familiar with the five logical connectives, you should start taking shortcuts. For instance, building a PMD for the wff ((Pv~Q)&Q)⇒P would normally take eight diagrams, but instead, ask yourself questions like the following -

i) If P is true and Q is true, what is (Pv~Q)?
ii) Now if (Pv~Q) is true and Q is true, what is ((Pv~Q)&Q)?
iii) If ((Pv~Q)&Q) is true and P is true, what is ((Pv~Q)&Q)⇒P?

By asking such questions, you should be able to fill in the PMD for ((Pv~Q)&Q)⇒P directly, without the need for any intermediate diagrams.

Normal Forms

We have seen that different wffs may be logically equivalent, and we are quite justified in using whichever form of a wff that we find most useful. For instance, since P⇒Q and ~PvQ and ~(P&~Q) are all equivalent, we may chose to use whichever one suits our purpose. Usually we chose the form which makes the wff easy to read for humans, but this may not be ideal for processing by computer.

> Definition: Normal forms
> 1. An atom is either a propositional variable or a negated propositional variable.
> 2. An elementary disjunction is some wff of the form (A_1vA_2v...A_n) where A_1,A_2,...A_n are all atoms.
> 3. An elementary conjunction is some wff of the form (A_1&A_2&...A_n) where A_1,A_2,...A_n are all atoms.

eg P, Q, ~Q, and ~R are all atoms
(Pv~Pv~QvQ) is an elementary disjunction
(~R&~R&R&P&Q) is an elementary conjunction

Notice that when writing these normal forms, we actually violate our definition of a wff since we allow conjunctions and disjunctions with more than two components. Technically we should write (A_1&(A_2&...(A_{n-1}&A_n)...) instead of (A_1&A_2&...A_n).

In the minimal case where n=1, a single atom also counts as both an elementary disjunction and an elementary conjunction.

> Definition: Normal forms
> 4. A wff is in conjunctive normal form (CNF) if it is a conjunction of elementary disjunctions. ie in the form (D_1&D_2&...D_n) where D_1,D_2,...D_n are all elementary disjunctions.
> 5. A wff is in disjunctive normal form (DNF) if it is a disjunction of elementary conjunctions. ie in the form (C_1vC_2v...C_n) where C_1,C_2,...C_n are all elementary conjunctions.

Given any wff it is possible to write an equivalent wff in conjunctive normal form and also possible to write another equivalent wff in disjunctive normal form. Not only is this possible, but there is a mechanical procedure for converting any wff into either of these desired forms. These two normal forms make the wff more difficult to read, but they have several advantages.

Consider the elementary disjunction (PvQv~P). Being a disjunction, such a wff will be true whenever *any* of its disjuncts are true. But notice that this wff contains both the atom P and also the atom ~P. One of these must always be true, and hence the wff overall will always be true. In fact this observation gives us a general rule - **an elementary disjunction is tautologous iff it includes some propositional variable and also the negation of that variable.**

The converse holds for elementary conjunctions. For instance, the wff (P&Q&~P) contains both P and ~P (one of which must be false), and so it will always be false. In general - **an elementary conjunction is inconsistent iff it includes some propositional variable and also the negation of that variable.**

Furthermore, a wff in CNF will be tautologous iff all of its elementary disjunctions are tautologous, ie iff all of its elementary disjunctions include some propositional variable and also the negation of that variable. Conversely, a wff in DNF will be inconsistent iff all of its elementary conjunctions are inconsistent, ie iff all of its elementary conjunctions include some propositional variable and also the negation of that variable.

Re-writing a wff in DNF and CNF make it possible to test the wff without needing to draw up either a truth table or a PMD. For simple wffs this is clearly not needed, but for a long wff with many propositional variables, using these normal forms is much quicker that PMDs or truth tables. A large proportion of computer-based logic relies on these normal forms.

<u>Exercises - Normal forms</u>

1. Specify whether the following wffs are in CNF or DNF or both or neither -
   - i) (PvQ)&(Pv~P)&(RvS)
   - ii) PvQ
   - iii) Pv(Q&~Q)
   - iv) (PvQ)&~(QvR)
   - v) ~R
   - vi) (P&Qv(R&(S&~P)))
2. Which of the following wffs are tautologous?
   - i) (PvQvR)&(Pv~Qv~R)&(~PvQ)
   - ii) (Pv~QvQ)&P&(Rv~Q)&(~PvP)
   - iii) (~PvQvP)&(Rv~SvTvPv~R)&(Qv~Q)
3. Which of the following wffs are inconsistent?
   - i) (Q&R&R)v~RvRv(P&~P)
   - ii) Pv~Pv(Q&R&S)
   - iii) (~P&Q&P)v(R&~S&T&P&~R)v(Q&~Q)

Entailment

An <u>argument</u> consists of a set of propositions (called its <u>premises</u> or <u>assumptions</u>) from which certain other propositions are said to <u>follow</u>. The final proposition in the argument is called the <u>conclusion</u>. The chief concern in the study of formal logic is the <u>validity</u> of an argument; that is, whether the conclusion really does follow from the premises.

<u>Definition: Entailment</u>

A∴B (read "A <u>entails</u> B") is called a <u>sequent</u>. It claims that there is a valid argument that starts with A as premise and ends with B as conclusion (ie B is a consequence of A).

More generally, the sequent $A_1, A_2, \ldots A_n$ ∴ B claims there is a valid argument starting with the set of premises $\{A_1, A_2, \ldots A_n\}$ and ending with the conclusion B.

Saying that A≡B is the same as saying that A and B are <u>inter-derivable</u> (that is A ∴ B and also B∴ A).

We will be following a very strict notational format in writing down arguments - strict in both the physical layout and the logical rules which allow us to move from one proposition to the next. The following is a fully worked example to give you a flavour of this notation.

Parliamentary example

> "If Parliament sits longer, more questions are raised. Either Parliament sits longer or the evasion of questions will widen the credibility gap. Now if it is true that the evasion of questions carries with it an increase in public relations effort, then it is also true that any widening of the credibility gap will make propaganda ineffective. Now certainly either propaganda is remaining effective or public relations efforts are not increasing. Coupling this with the known fact that the combination of Parliament not sitting longer and there being no increase in the number of questions raised will force any lessening of credibility to cause an increased effort in public relations, and bearing in mind that no more questions <u>are</u> being raised, it follows that either the credibility gap is not widened, or questions are not being evaded."

Analysing this in English is too ghastly to contemplate, but if we can codify each claim in propositional terms, we may have a better chance of proving or disproving the supposed conclusion.

To start with, we will ignore a lot of superfluous phrases (like "now certainly" and "coupling this with the known fact") and assign propositional variables to all the key phrases -

P = "Parliament sits longer"
Q = "more questions are raised"
E = "questions are being evaded"
C = "the credibility gap widens"
R = "public relations efforts increase"
I = "propaganda is ineffective"

We can then state each claim in the argument as a wff -

1. P⇒Q
2. Pv(E⇒C)
3. (E⇒R)⇒(C⇒I)
4. ~Iv~R
5. (~P&~Q)⇒(C⇒R)
6. ~Q

and the final conclusion as the wff -

7. ~Cv~E

We could put these together into one loooong wff in the form (((((1&2)&3)&4)&5)&6)⇒7 and then check whether this is a tautology by constructing a PMD (or truth table). But this would be daunting since the wff would be very long and the PMD would have 64 nodes. Instead, we want to see whether we can construct an argument, starting with the premises 1 to 6 and ending with the conclusion 7. That is, we want to piece together a series of deductions to show that propositions 1 to 6 entail 7. We write this assertion as the sequent 1,2,3,4,5,6 ∴ 7.

ie P⇒Q, Pv(E⇒C), (E⇒R)⇒(C⇒I), ~Iv~R, (~P&~Q)⇒(C⇒R), ~Q ∴ ~Cv~E

Now we need some procedure for writing down an argument in some formal way. This must include a set of rules which will allow each step of the argument to follow logically from previous steps. Such an argument, which starts with the six premises and finishes with the required conclusion, is shown below. Don't be too scared by the details of this formal argument - the main point is to see what is meant by logical entailment. After working through the sections which follow, you should be able to come back and fully understand this example.

1	(1)	P⇒Q	A
2	(2)	Pv(E⇒C)	A
3	(3)	(E⇒R)⇒(C⇒I)	A
4	(4)	~Iv~R	A
5	(5)	(~P&~Q)⇒(C⇒R)	A
6	(6)	~Q	A
1,6	(7)	~P	1,6 MTT
1,6	(8)	~P&~Q	6,7 &I
1,5,6	(9)	C⇒R	5,8 MPP
10	(10)	~(E⇒C)	A
1,6,10	(11)	~P&~(E⇒C)	7,10 &I
1,6,10	(12)	~(Pv(E⇒C))	11 SI (de Morgan)[1]
1,2,6,10	(13)	(Pv(E⇒C)) & ~(Pv(E⇒C))	2,12 &I
1,2,6	(14)	E⇒C	10,13 RAA
15	(15)	E	A
1,2,6,15	(16)	C	14,15 MPP
1,2,5,6,15	(17)	R	9,16 MPP
1,2,5,6	(18)	E⇒R	15,17 CP
1,2,3,5,6	(19)	C⇒I	3,18 MPP
20	(20)	~I	A
1,2,3,5,6,20	(21)	~C	19,20 MTT
1,2,3,5,6,20	(22)	~Cv~E	21 vI
23	(23)	~R	A
1,2,5,6,23	(24)	~E	18,23 MTT
1,2,5,6,23	(25)	~Cv~E	24 vI
1,2,3,4,5,6	(26)	~Cv~E	4, (20,22), (23,25) vE

Notice that the step by step deductions appear in the third column, alongside a line number. The right-hand column gives the reason for believing that line (one of the 10 Rules of Derivation described below). The left-hand column indicates the list of assumptions upon which each line depends. All of our proofs will have these four columns.

Having constructed an argument like the one above, we can happily say that we have proved the sequent. In other words, the English argument is valid; if premises 1 to 6 are true then the conclusion *must* also be true; the premises *do* entail the conclusion. If someone wants to deny the conclusion, they will not be able to do it on logical grounds; their only option will be to show that one or more of the premises are false.

Rules of Derivation

In order to refute an *informal* argument, various strategies may be used:
- claim the speaker has made an incorrect assumption
- claim the speaker has made an invalid deduction from the available evidence
- provide a counter-example
- questions the speaker's authority
- question the speaker's motives or integrity
- find photos of the speaker in the middle of a juicy sex scene etc etc

In *formal* logic we are primarily interested in the second of these strategies. The first and third may also interest us but the rest are irrelevant. In the light of the first two strategies it will be important to keep a clear distinction between propositions in an argument which have been assumed and those which have been deduced from the assumptions. Further, it will be important to keep track of which deductions are based on which assumptions.

---

[1] I've actually cheated a bit here by assuming de Morgan's law. To properly show that line 12 follows from line 11 would take about another 13 lines.

In the light of the second strategy, it would be useful if we could agree on a set of valid forms of deduction. If each step of the argument faithfully adheres to these agreed forms, then it will be difficult for anyone to deny the validity of the overall argument. The following 10 Rules of Derivation are defined with two aims -

i) to be simple and concise so that all the forms of deduction are obviously true to our intuitions

ii) though simple, the rules should be powerful enough so that any valid argument which can be expressed in propositional logic can be proved using these rules

1. Rule of Assumption (A)

> Definition: A
> Any proposition may be introduced at any stage of the proof as an assumption.

Initially, this seems like a very generous rule, but one should note that it is not the logicians duty to check out the truth of any of the assumptions: we are just concerned with what logically follows from an assumption given that it is true.

2. Modus Ponendo Ponens (MPP)

The Latin name modus ponendo ponens (sometimes refered to as simply modus ponens) means "the mode of reasoning which affirms something positive and deduces something positive".

> Definition: MPP
> Given A and A⇒B, we may derive B as conclusion. B depends on any assumptions on which either A or A⇒B depend. In other words, if a conditional holds and also its antecedent, then it's consequent also holds.

eg_1 P, P⇒Q ∴ Q is a valid sequent as proved by the following argument -

1	(1)	P⇒Q	A
2	(2)	P	A
1,2	(3)	Q	1,2 MPP

Column 1 shows the list of assumptions on which each line depends
Column 2 is simply a line number
Column 3 shows the wffs being assumed and deduced.
Column 4 is a justification which explains the reason for believing each line. This justification will includes the name of the Rule being used along with the line numbers of previous lines which are used as premises for that Rule.

Notice that the final line shows the conclusion which is the right-hand side of the sequent and that the assumption upon which the last line rest are exactly the wffs which are listed on the left-hand side of the sequent. This should always be the case; in fact that is what we mean by saying that the sequent is proved by the argument.

eg_2 P⇒Q, Q⇒R, P ∴ R

1	(1)	P⇒Q	A
2	(2)	Q⇒R	A
3	(3)	P	A
1,3	(4)	Q	1,3 MPP
1,2,3	(5)	R	2,4 MPP

Beware of the Fallacy of Affirming the Consequent. You may be tempted to represent the argument "Communists believe in redistribution of wealth, Mr X believes in redistribution of wealth, so he must be a Communist" by the sequent $C \Rightarrow R, R \therefore C$. But this is not a valid use of MPP.

3. Modus Tollendo Tollens (MTT)

The Latin name modus tollendo tollens (sometimes refered to as simply modus tollens) means "the mode of reasoning which starts with something negative and deduces something negative".

> Definition: MTT
> Given $\sim B$ and $A \Rightarrow B$, we may derive $\sim A$ as conclusion. $\sim A$ depends on any assumptions on which either $\sim B$ or $A \Rightarrow B$ depend. In other words, if a conditional holds and its consequent is denied then the antecedent can also be denied.

eg$_1$ "If de Klerk were the President of the USA then he would have to be a US citizen. But de Klerk is not a US citizen, hence he cannot be the US President."

Using the dictionary - P = "de Klerk is President of the USA"
C = "de Klerk is a US citizen"

we can represent this argument as the sequent – $P \Rightarrow C, \sim C \therefore \sim P$
and can prove it to be valid as follows -

1	(1)	$P \Rightarrow C$	A
2	(2)	$\sim C$	A
1,2	(3)	$\sim P$	1,2 MTT

eg$_2$ $P \Rightarrow (Q \Rightarrow R), P, \sim R \therefore \sim Q$

1	(1)	$P \Rightarrow (Q \Rightarrow R)$	A
2	(2)	P	A
3	(3)	$\sim R$	A
1,2	(4)	$Q \Rightarrow R$	1,2 MPP
1,2,3	(5)	$\sim Q$	4,3 MTT

Beware of the Fallacy of Denying the Antecedent. You may be tempted to represent the argument "If the program has a syntax error then it will not give the correct solution. The program has no syntax errors. Therefore, the program will give the correct solution." by the sequent $S \Rightarrow \sim C, \sim S \therefore C$. But this is not a valid use of MTT.

4. Double Negation (DN)

> Definition: DN
> Given A we may derive $\sim\sim A$ and vice-versa. The conclusion depends on the same assumptions as the premise.

eg$_1$ $P \therefore \sim\sim P$, by the following proof -

1	(1)	P	A
1	(2)	$\sim\sim P$	1 DN

eg2 ~P⇒Q, ~Q ∴ P

1	(1)	~P⇒Q	A
2	(2)	~Q	A
1,2	(3)	~~P	1,2 MTT
1,2	(4)	P	3 DN

5. Conditional Proof (CP)

Definition: CP
Given a proof of B from A as assumption, we may derive A⇒B as conclusion based on the remaining assumptions (if any).

In other words, if you could prove P ∴ Q then you could equally well prove ∴ P⇒Q. CP is one of only two Rules which reduces the number of assumptions (RAA is the other one).

eg P⇒Q ∴ ~Q⇒~P

1	(1)	P⇒Q	A
2	(2)	~Q	A
1,2	(3)	~P	1,2 MTT
1	(4)	~Q⇒~P	2,3 CP

Exercises: Proofs
Construct proofs for each of the following sequents -
i) P⇒~Q, Q ∴ ~P
ii) ~Q⇒~P ∴ P⇒Q
iii) P⇒Q, R⇒~Q ∴ P⇒~R

6. &-Introduction (&I)

Definition: &I
Given A and B separately, we may derive A&B as conclusion. The conclusion rests on all assumptions on which either A or B depend.

eg1 P,Q ∴ P&Q

1	(1)	P	A
2	(2)	Q	A
1,2	(3)	P&Q	1,2 &I

7. &-Elimination (&E)

Definition: &E
Given A&B, we may derive either A or B as conclusion. The conclusion depends on the same assumptions as the premise.

eg P&Q ∴ Q&P

1	(1)	P&Q	A
1	(2)	P	1 &E
1	(3)	Q	1 &E
1	(3)	Q&P	3,2&I

<u>Exercises: Proofs</u>
Construct proofs for each of the following sequents -

i) P, P⇔Q ∴ Q
ii) P&(Q&R) ∴ (P&Q)&R
iii) (P&Q)⇒R ∴ P⇒(Q⇒R)
iv) P⇒(Q⇒R) ∴ (P&Q)⇒R

(Note that proving sequent iii is not the same as proving sequent iv; both must be proved separately.)

8. v-Introduction (vI)

<u>Definition: vI</u>
Given either A or B separately, we may derive AvB as conclusion. The conclusion depends on the same assumptions as the premise.

eg "Charles I was beheaded. Therefore, either Charles I was beheaded or he was sent to the electric chair." - a very dull conclusion, but nevertheless quite valid.

9. v-Elimination (vE)

<u>Definition: vE</u>
Given AvB, together with a proof of C from A as assumption and a proof of C from B as assumption, we may derive C as conclusion. C depends on any assumption on which AvB depends, or on which C depends in its derivation from A (apart from A), or on which C depends in its derivation from B (apart from B).

In other words, if you can prove P ∴ R and can also prove Q ∴ R then you have also proved PvQ ∴ R. This is the most confusing of the 10 Rules. Proofs using vE will always have the following structure (where <Ass1>,<Ass2>, and <Ass3> represent lists of auxiliary assumptions) -

	.	
	.	
<Ass1>	(10) AvB	?
11	(11) A	A
	.	
11, <Ass2>	(19) C	?
20	(20) B	A
	.	
20, <Ass3>	(29) C	?
<Ass1>,<Ass2>,<Ass3>	(30) C	10,(11,19),(20,29) vE

Notice that the conclusion C is written three times: once as a deduction from A (and possibly some other assumptions), once as a deduction from B (and possibly some other assumptions), and then the final time in the vE step. This will always be the case. It would be wrong to stop the proof at line 19, because although we have reached the required conclusion, it is not based on the required assumptions.

The lines quoted as justification for vE are firstly the original disjunction (10), secondly the first subproof from A to C (11 to 19), and thirdly the second subproof from B to C (20 to 29).

eg PvQ ∴ QvP

1	(1)	PvQ	A
2	(2)	P	A
2	(3)	QvP	2 vI
4	(4)	Q	A
4	(5)	QvP	4 vI
1	(6)	QvP	1,(2,3),(4,5) vE

<u>Exercise: vE Proof</u>
1. Construct a proof for the sequent Q⇒R ∴ (PvQ)⇒(PvR)
2. Construct a proof for the sequent (P⇒Q)&(R⇒S), PvR ∴ QvS

10. Reductio ad Absurdum (RAA)

If you ever reach a contradiction, you may be quite sure that at least one of your assumptions is false. From now on we will use a more exact definition of "contradiction" than previously. A <u>contradiction</u> is any wff of the form α&~α; that is, a conjunction in which the second conjunct is the negation of the first. Thus we may describe the Rule whose Latin name means "to reduce to absurdity" -

<u>Definition: RAA</u>
Given a proof of (B&~B) from A as assumption, we may derive ~A as conclusion based on the remaining assumptions (if any).

eg P⇒Q, P⇒~Q ∴ ~P

1	(1)	P⇒Q	A
2	(2)	P⇒~Q	A
3	(3)	P	A
1,3	(4)	Q	1,3 MPP
2,3	(5)	~Q	2,3 MPP
1,2,3	(6)	Q&~Q	4,5 &I
1,2	(7)	~P	3,6 RAA

The contradiction at line 6 shows that assumptions 1,2 and 3 are inconsistent: they cannot all be true. RAA allows us to reject *any one* of those assumptions: in this case we have chosen to reject 3, but we could have written either of the following as line 7 -

2,3	(7)	~(P⇒Q)	1,6 RAA
1,3	(7)	~(P⇒~Q)	2,6 RAA

<u>Exercises: RAA Proofs</u>
Construct proofs for the sequents -
i) P⇒Q ∴ (P⇒~Q)⇒~P
ii) P, ~(P&Q) ∴ ~Q

Other forms of deduction

There are many other common forms of valid deduction, but these can all be proved using the 10 rules above. For instance -

Hypothetical Syllogism	P⇒Q, Q⇒R ∴ P⇒R
Constructive Dilemma	(P⇒Q)&(R⇒S), PvR ∴ QvS
Destructive Dilemma	(P⇒Q)&(R⇒S), ~Qv~S ∴ ~Pv~R
Modus Ponendo Tollens	~(P&Q), P ∴ ~Q
Modus Tollendo Ponens	PvQ, ~P ∴ Q
Resolution	PvQ, ~PvR ∴ QvR

Exercises

1. Examine the proof of Russell's Paradox. On which of the Rules of Derivation does it rely?
2. Consider the following argument -

   "University students will either study or they will party. If they study they are assured of passing and once they have passed they will graduate. However, if they party they will surely fail. Those students who fail eventually drop out of university. So it is clear that all students eventually leave university."

   i) List the Rules of Derivation which are used in the argument.
   ii) Is the argument valid?
   iii) Is it sound?
   iv) Create a dictionary to represent the propositions in the argument.
   v) Express the argument as a series of propositional wffs (you may also need to introduce an extra premise which is assumed but unstated in the English argument).
   vi) Construct a formal proof which mirrors the English argument.
3. Review the parliamentary argument on page 39 and check that each line uses the Rules of Derivation correctly.

Sequent Introduction (SI)

Proving sequents using only the 10 allowed Rules is tedious work and in order to save time, brain cells and ink, I now introduce a short cut. Suppose half way through a proof we want to use one of the sequents which we have already proved. Of course we could copy each line from the previous proof (in effect reproving the result), but it is more sensible to simply cite the other proof and carry on. Citing a previous proof is called sequent introduction and must conform to the following pattern -

Definition: SI

If $A_1, A_2, \ldots A_n \therefore$ B is a proven sequent and the wffs $A_1, A_2, \ldots A_n$ have all been established, then we may derive B as conclusion. The assumptions on which B depends will be all the assumptions on which $A_1, A_2, \ldots A_n$ depend.

For example, suppose we want to prove the principle of modus tollendo ponens. Although this looks like a fairly straight-forward sequent, a full proof is actually quite lengthy, and so we will do the proof in three parts.

1. We first prove the sequent PvQ ∴ ~(~P&~Q)

1	(1)	PvQ	A	
2	(2)	P	A	Start of 1st vE subproof
3	(3)	~P&~Q	A	Start of RAA
3	(4)	~P	3 &E	
2,3	(5)	P&~P	2,4 &I	
2	(6)	~(~P&~Q)	3,5 RAA	End of 1st vE subproof
7	(7)	Q	A	Start of 2nd vE subproof
3	(8)	~Q	3 &E	
3,7	(9)	Q&~Q	7,8 &I	
7	(10)	~(~P&~Q)	7,9 RAA	End of 2nd vE subproof
1	(11)	~(~P&~Q)	1,(2,6),(7,10) vE	

2. Then we prove ~(~P&~Q) ∴ ~P⇒Q

1	(1)	~(~P&~Q)	A	
2	(2)	~P	A	Start of CP
3	(3)	~Q	A	Start of RAA
2,3	(4)	~P&~Q	2,3 &I	
1,2,3	(5)	(~P&~Q)&~(~P&~Q)	4,1 &I	
1,2	(6)	~~Q	3,5 RAA	
1,2	(7)	Q	6 DN	
1	(8)	~P⇒Q	2,7 CP	

3. If we make use of these two previously proved sequents then the final proof of PvQ, ~P ∴ Q becomes short -

1	(1)	PvQ	A
1	(2)	~(~P&~Q)	1 SI PvQ ∴ ~(~P&~Q)
1	(3)	~P⇒Q	2 SI ~(~P&~Q) ∴ ~P⇒Q
4	(4)	~P	A
1,4	(5)	Q	3,4 MPP

This simple form of sequent introduction is useful, but what happens if you have proved the sequent P ∴ Q⇒ P but in the proof you are working on you require (A&B) ∴ Q ⇒ (A&B)? The two have the same logical form and hence a proof for the first can easily be turned into a proof for the second by replacing every occurrence of P with (A&B). So let's extend Sequent Introduction -

> <u>Definition: Sequent Introduction with Substitution</u>
> If $A_1,A_2,...A_n$ ∴ B is a proven sequent and $A'_1,A'_2,...A'_n$ ∴ B' is a <u>substitution instance</u> of $A_1,A_2,...A_n$ ∴ B, and $A'_1,A'_2,...A'_n$ have all been established, then we may derive B' as conclusion. The assumptions on which B' depend are all the assumptions on which $A'_1,A'_2,...A'_n$ depend.

Meta-Theorems about Derivations

1. If $A_1,A_2,...A_n$∴B is a derivable sequent then so is $A_1,A_2,...A_{n-1}$∴A_n⇒B.
<u>Proof:</u> The assumption that $A_1,A_2,...A_n$∴B is a derivable sequent means that there is some argument in the form -

1	(1)	A_1	A
2	(2)	A_2	A
	.		
	.		
n	(n)	A_n	A
	.		
	.		
1,2,...n	(x)	B	???

Given such a proof, we can always add the next line -
1,2,...n-1 (x+1) A_n⇒B n,x CP
which constitutes a proof of $A_1,A_2,...A_{n-1}$∴A_n⇒B.

2. The converse of 1. is also true: If $A_1,A_2,...A_{n-1}$∴A_n⇒B is a derivable sequent then so is $A_1,A_2,...A_n$∴B.

Proof: The assumption that $A_1, A_2, \ldots A_{n-1} \therefore A_n \Rightarrow B$ is a derivable sequent means that there is some argument in the form -

1	(1)	A_1	A
2	(2)	A_2	A
	.		
	.		
n-1	(n-1)	A_{n-1}	A
	.		
	.		
1,2,...n-1	(x)	$A_n \Rightarrow B$	???

Given such a proof, we can always add the two lines -

x+1	(x+1)	A_n	A
1,2,...n-1,x+1	(x+2)	B	x, x+1 MPP

which constitutes a proof of $A_1, A_2, \ldots A_n \therefore B$.

3. By repeated applications of 1. and 2. one can see that $A_1, A_2, \ldots A_n \therefore B$ can be proved iff $\therefore (A_1 \Rightarrow (A_2 \Rightarrow (A_3 \ldots (A_n \Rightarrow B) \ldots)$ can be proved.

4. $\therefore$ A iff A is a tautology.

5. Proving $A_1, A_2, \ldots A_n \therefore B$ is precisely the same as showing that $(A_1 \Rightarrow (A_2 \Rightarrow (A_3 \ldots (A_n \Rightarrow B) \ldots)$ is a tautology. (By combining 3. and 4.)

Comparison of Arguments, Truth Tables and Possible Models Diagrams

1. The sequent $A_1, A_2, \ldots A_n \therefore B$ may be proved in three ways -
   i) construct an argument which assumes $A_1, A_2, \ldots A_n$ and deduces B using the 10 Rules of Derivation;
   ii) show that $(A_1 \Rightarrow (A_2 \Rightarrow (A_3 \ldots (A_n \Rightarrow B) \ldots)$ is a tautology (either by truth table or PMD);
   iii) show that the PMD for $(A_1 \& (A_2 \& (A_3 \ldots (A_{n-1} \& A_n) \ldots)$ is a subset of the PMD for B.

Exercises: Alternate methods of proof

1. We have already constructed formal arguments which prove the following sequents -
   i) $Q \Rightarrow R \therefore (PvQ) \Rightarrow (PvR)$
   ii) $P \Rightarrow Q, P \Rightarrow \sim Q \therefore \sim P$
   iii) $P \Rightarrow Q \therefore (P \Rightarrow \sim Q) \Rightarrow \sim P$
   iv) $P, \sim(P \& Q) \therefore \sim Q$

   Now re-prove each of these by the other two methods.
2. Show that the following arguments are invalid -
   i) If Alice wins first prize, then Bob wins second prize and if Bob wins second prize then Carol is disappointed. Either Alice wins first prize or Carol is disappointed. Therefore Bob does not win second prize.
   ii) If the seed catalogue is correct, then if the seeds are planted in April, then the flowers bloom in July. The flowers did bloom in July. Therefore, if the seeds catalogue is correct, then the seeds must have been planted in April.

2. An argument may prove a sequent but cannot be used to disprove it. Truth tables and PMDs may be used for both.

3. Truth tables and PMDs are easy to handle if there is a small number of propositional variables, but it is generally easier to construct an argument if there are more than three.

4. Truth tables and PMDs are only useful for propositional logic, whereas the argument structure and 10 Rules of Derivation work for both propositional and predicate logic.

5. PMDs have more visual appeal that truth tables, but truth tables are more standard in computer science.

Consistency and Completeness

There are two important questions to ask about our 10 Rules of Derivation:
1. Are the Rules safe? ie can we guarantee that all the sequents we can prove are reliable?
2. Are the Rules powerful enough, or might there be some valid sequents which we cannot prove?

More formally, we need to judge this system of logic on the basis of the following two criteria -

> Definition: Consistency
> If every sequent provable by a system of logic is necessarily true (ie tautologous), then the system is said to be consistent.

> Definition: Completeness
> If every necessary truth which can be represented in a system of logic can be proved to be valid in that system, then the system is said to be complete.

Propositional logic based on the 10 Rules of Derivation form a system which is both consistent and complete.[1]

Axiomatic Basis for Propositional Logic

There is always a trade off between the number of Rules of Derivation and the complexity of proofs. If one allows more Rules, then proofs become shorter; but the more Rules there are, the more difficult it becomes to be sure that the system is complete and consistent. Our system of 10 Rules is a compromise: bigger than necessary, but not too big to become cumbersome.

Bertrand Russell and Alfred North Whitehead defined a much smaller system which still encapsulates everything that our system can do. This system can be defined axiomatically as follows -

1. We allow any letter to be used as a propositional variable; the symbols (,),v, &, ~ and ⇒ (all defined as expected); and the formation of wffs as previously defined.
2. The following wffs are axioms -
   - A1: (PvQ)⇒P
   - A2: Q⇒(PvQ)
   - A3: (PvQ)⇒(QvP)
   - A4: (Q⇒R)⇒((PvQ)⇒(PvR))
3. The Rule of Uniform Substitution: if you select any variable in a theorem and replace every occurrence of that variable with any wff, then the resulting wff is also a theorem.
4. Modus Ponens: if α and β represent any two wffs and $(\alpha \Rightarrow \beta)$ and α are both theorems then so is β.

---

[1] These two concepts are also important guides to judging computing systems (eg evaluating a CPU chip design, or a new hardware architecture, or a new programming language).
- Completeness: will the system be able to compute everything which is computable? (A question which is answered in the CS2 Theory of Computing module.)
- Consistency: can you guarantee that the system will be error-free?

Section 6 - Predicate Logic

Inadequacies of Propositional Logic

Propositional logic treats all propositions as atomic (ie indivisible, without any internal structure). For many statements and arguments this is inadequate. For instance it seems like an oversimplification to represent the sentence "Every person in this room is awake" by just a single propositional variable P. We could write a long wff such as P&(Q&(R&(S&(......))) to represent "Peter is awake and Themba is awake and Sue is awake ..." but that still doesn't seem to capture the meaning of the original statement. So an argument like "Every person in this room is awake. Matthew Clarke is in this room. Therefore Matthew Clarke is awake.", which is clearly valid, is not provable (nor even expressible) in propositional logic.

To overcome such shortcomings we introduce new notation to express properties and relations, notation to represent variables, notation to represent the quantities "all" and "some", and finally four new Rules of Derivation.

Predicates

We introduce predicates to represent claims that the subject of a proposition has some property (or characteristic). For instance, suppose we want to represent the claims -

1. "John is a bachelor"; and
2. "Peter is a bachelor".

If we let B be a predicate symbolising "bachelor" we can rewrite 1. as B(John) and 2. as B(Peter).

We can also say that for any $x \in$ {John, Peter}, it is true that B(x). The x here is a variable which does not evaluate to true or false, but merely represents an object. Since x may represent any number of things, we don't even want to say that B(x) is either true or false. However, when we consider an actual object, say John, then B(John) must be either true or false.

"Bachelor" is an example of a one-place predicate (representing a property of an object), but we can also make use of multi-place predicates (representing relationships between objects). For instance, if we let j_1 represent Jack, j_2 represent Jill, and let T be the predicate "taller than", then $T(j_1,j_2)$ means "Jack is taller than Jill". Note that the order of objects is important in multi-place predicates: $T(j_1,j_2)$is not the same as $T(j_2,j_1)$.

In propositional logic we could use upper and lower case letters interchangeably, but now we must be more careful. We will always use capitals to represent predicates; x,y,z as variables; and other lower case letters (normally a,b,c) as proper names representing particular objects.

Quantification

Universal Quantification

Suppose we wish to represent the proposition "all people are mad". Letting M be the predicate "mad", we could write M(x), but this is simply a claim about some object x. To apply this to all possible x's, we write ($\forall$ x)M(x). (In some logics, this may just be written (x)Mx[1].)

---

1 The notation (x)Mx is used by Lemmon for instance (and hence in the LemmonAid software). Although it makes wffs longer and the proliferation of brackets may be confusing, I have chosen to use the more verbose notation ($\forall$ x)M(x) for two reasons. Firstly, it is good to be familiar with the $\forall$ symbol because it is so widely used. Second, using extra brackets complies with standard computing notation (eg Prolog) where the predicates are seen as functions and the variables as parameters.

Definition: Universal Quantifier
The symbol "∀" is the universal quantifier and is read "for all".

Thus, (∀x)M(x) can be read as any of the following -

For every x in the universe, x is mad
For all x, x is mad
Everything is mad

If we want to make this a bit less universal, so that it applies only to humans, we could introduce another predicate H for "human" and write (∀x)(H(x)⇒M(x)). This can be read "for all x, if x is human then x is mad", or in short "every human is mad". Note that this is not the same as the wff (∀x)(H(x)&M(x)) - this claims that everything in the universe is both human and mad.

As another example, we could symbolise the fact that being taller than someone is non-reflexive by the expression (∀x)(∀y)(T(x,y)⇒~T(y,x)).

To symbolically represent "All romans were either loyal to Caesar or hated him" we could write -

(∀x)(Roman(x) ⇒ (loyal_to(x,Caesar) v hates(x,Caesar)))

though we will more typically use only single letters and write -

(∀x)(R(x) ⇒ (L(x,c) v H(x,c)))

It is often useful to think of a universally quantified wff as an *infinite conjunction*. For instance if all the elements of the universal set were labelled $a_1,a_2,a_3,...$ then (∀x)P(x) is equivalent to $P(a_1)\&P(a_2)\&P(a_3)\&$ ...

Exercise: Universal quantification
Express the sentence "Every person in this room is awake" as a predicate wff.

Existential Quantification

Using universal quantification we could represent claims such as "all unicorns are alcoholics", but this seems to miss out on the important fact that, unfortunately, there are no unicorns. To rectify this we introduce the existential quantifier.

Definition: Existential Quantifier
The symbol "∃" is the existential quantifier and is read "there exists".

eg If A is a predicate meaning "asleep in class" then (∃x)A(x) means "there is at least one x such that x is asleep in class", or more loosely "something (somebody?) is asleep in class".

It is often useful to think of an existentially quantified wff as an *infinite disjunction*. For instance if all the elements of the universal set were labelled $a_1,a_2,a_3,...$ then (∃x)P(x) is equivalent to $P(a_1)vP(a_2)vP(a_3)v$ ...

Exercises: Existential quantification
1. Using the same predicates as the previous exercise, express the sentence "At least one person in this room is awake" as a predicate wff.
2. Express the proposition "There is at least one honest politician" as a predicate wff.
3. Express the proposition "There are no unicorns" as a predicate wff.

Exercises: Representing English in predicate notation

For each of the following sentences, devise a dictionary of predicates and then express the sentence as a wff of predicate logic -

- i) People are either male or female
- ii) No woman is both a doctor and a school student
- iii) Only men are whisky drinkers
- iv) All birds fly
- v) No bird flies
- vi) Not all birds fly
- vii) Tweety is a bird who doesn't fly
- viii) Every bit is either on or off
- ix) Either every bit is on or every bit is off
- x) Some computers are faster than others
- xi) A garfinkel cannot be snerked, but if it is transfixed, it snerks all its transfixers.

Free and Bound Variables

Consider -

1. $S(x) \Rightarrow (\exists y)T(x,y)$
2. $(\exists x)(P(x)\ \&\ (\exists y)T(x,y))$
3. $(\exists x)P(x)\ \&\ (\exists y)T(x,y)$

In 1. only the variable y has a quantifier. We say that y is a <u>bound variable</u> and that x is a <u>free variable</u>. We could interpret this expression in one and only one way, regardless of whether we used parentheses to indicate which y-values were quantified by (∃y).

However, in 2. the parenthesis following (∃y) is absolutely essential: 2. is <u>not</u> the same as 3. Without the parenthesis we are not saying that the x satisfying P(x) is necessarily the same x that satisfies T(x,y). In 3. the first x is bound, but the second is free.

In an expression (∀x)(...) or (∃x)(...), the portion of the expression to which the (∀x) or (∃x) applies is called the <u>scope</u> of the quantifier. Scope is indicated by parentheses except where it is obvious anyway. A variable is said to be <u>bound</u> by a quantifier which specifies that same variable if it lies within the scope of that quantifier. A variable which is not bound by any quantifier is said to be <u>free</u>.

The concept of free and bound variables should be familiar to you from maths. In the formula -

$$\int_1^2 3x^2y.dx$$

you may set y to any value you like and then evaluate the integral by varying x from 1 to 2. The variable x is <u>bound</u> by the integration operator, but y is <u>free</u> to be set outside the scope of the integral.

The terms free and bound are quite common in various areas of computer science as well as in many logic textbooks, but in <u>this</u> presentation of predicate logic we tend to avoid the confusion by using proper names (a,b,c) when we want free variables and variables (x,y,z) when we want bound variables. In terms more appropriate to computer programming, a proper name parallels the concept of a Pascal *constant* while a variable parallels a Pascal *variable*.

<u>Exercises: Free and bound variables</u>
Write down each variable in the following wffs and state whether they are free or bound. Also indicate the scope of each quantifier -

i) $(\forall x)G(x,y)$
ii) $(\forall x)(\exists y)(G(x,y) v P(x))$
iii) $(\forall x)(\exists y)G(x,y) v P(x)$
iv) $(\forall y)G(x,y) v (\exists y)P(x)$
v) $(\exists x)(H(x) \Rightarrow ((\forall y)K(y,x) \& H(a))$

Alternate Representations of English Statements

In many cases an English statement may be represented logically in more than one way. For instance, "No man is an island unto himself" may be represented by either of the following wffs (and certainly others as well) -

i) $\sim(\exists x)(M(x) \& I(x))$
ii) $(\forall x)(M(x) \Rightarrow \sim I(x))$

This can occur for one of two reasons: the alternate wffs may be *logically equivalent* or the English statement may be *ambiguous.*

1. Equivalence

As defined already in the section on propositional logic, two wffs are equivalent iff they are true under exactly the same circumstances. In propositional logic we could test this by constructing either a truth table or a PMD. That can also be of some use with predicate wffs, but truth tables cannot take account of quantification.

Apart from the equivalences we have already examined, there are only four extra ones which need to be learnt. These equivalences allow us to interchange universal and existential quantifiers -

i) $(\forall x)W(x) \equiv \sim(\exists x)\sim W(x)$
ii) $(\forall x)\sim W(x) \equiv \sim(\exists x)W(x)$
iii) $(\exists x)W(x) \equiv \sim(\forall x)\sim W(x)$
iv) $(\exists x)\sim W(x) \equiv \sim(\forall x)W(x)$

Each of these make intuitive sense (and later on we will be able to prove them conclusively). For instance the second equivalence shows that "everything in the universe is not a unicorn" is exactly the same as "it is not the case that there exists a unicorn". In each case, W(x) need not be simply a predicate, but could be any expression which contains the variable x.

eg Using the first equivalence, $(\forall x)(H(x) \Rightarrow (\forall y)L(x,y))$ is logically equivalent to $\sim(\exists x)\sim(H(x) \Rightarrow (\forall y)L(x,y))$

Furthermore, if the two expressions $W_1(x)$ and $W_2(x)$ can be shown to be equivalent, then $(\forall x)W_1(x) \equiv (\forall x)W_2(x)$ and $(\exists x)W_1(x) \equiv (\exists x)W_2(x)$.

eg if $W_1(x)$ is $\sim(P(x) \& Q(x))$ and $W_2(x)$ is $\sim P(x) v \sim Q(x)$,
we can show that $W_1(x) \equiv W_2(x)$ by a truth table or PMD test
and hence can conclude that $(\forall x)\sim(P(x) \& Q(x)) \equiv (\forall x)(\sim P(x) v \sim Q(x))$.

2. Ambiguity

All natural languages (such as English) contain ambiguities. For instance, the sentence "I love her cooking" can be interpreted in at least three separate ways (some people can see six ways).

As a simpler example, consider the sentence "Every boy loves a certain girl". This could mean that there is one (very fortunate) girl who is loved by every boy, in which case we could use the wff -

$$(\exists x)(G(x)\&(\forall y)(B(y)\Rightarrow L(y,x)))$$

But maybe the sentence is supposed to mean that for every boy you can find <u>some</u> girl whom he loves -

$$(\forall y)(B(y)\Rightarrow(\exists x)(G(x)\&L(y,x)))$$

These two wffs are not logically equivalent. They reflect the fact that the original English sentence is ambiguous: it is open to different understandings.

> <u>Exercise: Alternate representations</u>
> At the beginning of this section we proposed that "No man is an island unto himself" may be represented by either of the following wffs -
> i) $\sim(\exists x)(M(x)\&I(x))$
> ii) $(\forall x)(M(x)\Rightarrow\sim I(x))$
> Are these two wffs logically equivalent or is the original sentence ambiguous?

Rules of Derivation

All 10 Rules of Derivation from the propositional calculus carry over into the predicate calculus without modification. In addition, there are four rules governing the use of quantifiers.

11. Universal Elimination (UE)

If we know that for every object in the universe, some property holds, then for any particular object, that property must hold. eg Since all triangles have an angle sum of 180°, any actual triangle you care to chose will have an angle sum of 180°.

> <u>Definition: UE</u>
> Let x be a bound variable, F(x) be some expression (possibly just a single predicate but maybe a combination of predicates, propositional variables and connectives) and a be a free variable. Then given $(\forall x)F(x)$ we may derive F(a) as conclusion by replacing all occurrences of x in F(x) with a. The conclusion depends on the same assumptions as the premise.

eg Pascal is a programming language. All programming languages have loop structures. Therefore Pascal has a loop structure.

Dictionary - P(x) means x is a programming language
L(x) means x has a loop structure
s stands for Pascal

Then we need to prove the sequent $P(s), (\forall x)(P(x)\Rightarrow L(x)) \therefore L(s)$

1	(1)	$P(s)$	A
2	(2)	$(\forall x)(P(x)\Rightarrow L(x))$	A
2	(3)	$P(s)\Rightarrow L(s)$	2 UE
1,2	(4)	$L(s)$	3,1 MPP

If you think of a universally quantified expression as an infinite conjunction, then this rule is similar to &E. From P(a)&P(b)&P(c)& ... we can deduce any one of the conjuncts, say P(c).

Beware of negated quantifiers. From $(\forall x)\sim P(x)$ you can derive ~P(a) by UE, but from $\sim(\forall x)P(x)$ it is <u>not legal</u> to derive ~P(a). (This would be the same as deducing ~Q from the premise ~(Q&R).) Predicate wffs with negation signs at the front are difficult to handle (they normally require RAA proofs) and we generally try to avoid them.

12. Universal Introduction (UI)

By comparison with &I, if we could be sure that P(a), P(b), P(c) etc were all true (ie if P holds for each object in the universe) then we could deduce $(\forall x)P(x)$. Unfortunately this is rarely possible since the number of objects in the universe is infinite!

But compare this situation to proofs in geometry - suppose we wanted to prove that all triangles had an angle sum of 180°. Rather than try to prove this for each triangle in the universe, we would say "let ABC be any triangle, then ...". That is, we pick an arbitrary triangle which represents all triangles, prove the result for that triangle and deduce that it must be true for all triangles.

> Definition: UI
> Let F(a) be a wff containing the free variable a; let x be a bound variable not occurring in F(a); and let F(x) be formed by replacing all occurrences of a in F(a) with x. Then given F(a), we may derive $(\forall x)F(x)$ as conclusion, provided that a does not occur in any assumption on which F(a) rests. The conclusion depends on the same assumptions as the premise.

eg $(\forall x)(H(x)\Rightarrow G(x)) \therefore (\forall x)H(x)\Rightarrow(\forall y)G(y)$

1	(1)	$(\forall x)(H(x)\Rightarrow G(x))$	A
1	(2)	$H(a)\Rightarrow G(a)$	1 UE
3	(3)	$(\forall x)H(x)$	A
3	(4)	$H(a)$	3 UE
1,3	(5)	$G(a)$	2,4 MPP
1,3	(6)	$(\forall y)G(y)$	5 UI
1	(7)	$(\forall x)H(x)\Rightarrow(\forall y)G(y)$	3,6 CP

In summary, if an arbitrary proper name (in this case "a") can be shown to have some property, then UI allows us to conclude that everything in the universe has that property. We must be very careful about our choice of the proper name in order to guarantee that it is really arbitrary. **If "a" occurs in any of the assumptions upon which F(a) depends, then we cannot conclude $(\forall x)F(x)$.**

We must also take care about which bound variable to use. Having derived the wff $F(a)\Rightarrow(\forall x)G(x)$, we cannot use UI to deduce $(\forall x)(F(x)\Rightarrow(\forall x)G(x))$ - this isn't even a wff since x is quantified twice within the same scope. We could, however, deduce $(\forall y)(F(y)\Rightarrow(\forall x)G(x))$.

If these restrictions where not imposed, then we could prove the following -

1	(1)	$F(a)$	A
1	(2)	$(\forall x)F(x)$	1 UI

This sequent is plainly ludicrous. It is equivalent to the argument "Let a be an arbitrary number which is odd. Therefore all numbers are odd."

13. Existential Introduction (EI)

If we know that F(a) is true for any one object a (it doesn't even have to be an arbitrary object) then we can certainly conclude F(a)vF(b)vF(c)... by repeated uses of vI. Similarly we could conclude $(\exists x)F(x)$.

> **Definition: EI**
> Let x be a bound variable, F(x) be some expression (possibly just a single predicate but maybe a combination of predicates, propositional variables and connectives) and a be a free variable. Then given F(a) we may derive $(\exists x)F(x)$ as conclusion by replacing all occurrences of a in F(a) with x. The conclusion depends on the same assumptions as the premise.

eg $(\forall x)F(x) \therefore (\exists x)F(x)$

1	(1)	$(\forall x)F(x)$	A
1	(2)	$F(a)$	1 UE
1	(3)	$(\exists x)F(x)$	2 EI

14. Existential Elimination (EE)

If something has a certain property, and if it can be shown that some conclusion C follows from the assumption that an arbitrary object has that property, then we know that C must hold. In other words, if something has the property F, and no matter which has it then C holds, then C holds anyway.

> **Definition: EE**
> Let F(a) be a wff containing the free variable a; let x be a bound variable not occurring in F(a); and let F(x) be formed by replacing all occurrences of a in F(a) with x. Then given $(\exists x)F(x)$ together with a proof of some wff C from F(a) as assumption, we may derive C as conclusion, provided that a does not occur in C or in any assumption used to derive C from F(a) (apart from F(a) itself). The conclusion depends on any assumptions on which $(\exists x)F(x)$ depends or which are used to derive C from F(a) (apart from F(a)).

This rule is similar to vE. With vE we start with a disjunction AvB; assume A and prove C; assume B and prove C; then conclude C. With EE we start with an infinite disjunction, and rather than have an infinite number of subproofs (one for each disjunct), we just do one subproof using a typical disjunct. An EE proof will always have the following structure (where <Ass1> and <Ass2> represent lists of auxiliary assumptions)-

	.			
	.			
<Ass1>	(10)	$(\exists x)F(x)$	?	
11	(11)	$F(a)$	A	The typical disjunct
	.			
	.			
11, <Ass2>	(19)	C	?	
<Ass1>,<Ass2>	(20)	C	10,(11,19) EE	

eg $(\forall x)(F(x)\Rightarrow G(x)), (\exists x)F(x) \therefore (\exists x)G(x)$

1	(1)	$(\forall x)(F(x)\Rightarrow G(x))$	A
2	(2)	$(\exists x)F(x)$	A
1	(3)	$F(a)\Rightarrow G(a)$	1 UE
4	(4)	$F(a)$	A
1,4	(5)	$G(a)$	3,4 MPP
1,4	(6)	$(\exists x)G(x)$	5 EI
1,2	(7)	$(\exists x)G(x)$	2, (4,6) EE

Beware of the following restriction: in order to apply EE to remove an arbitrary name "a", the "a" must not appear in either the conclusion C or in any assumption used to conclude C (apart from the typical disjunct of course).

<u>Exercise: Restrictions on proofs using quantifiers</u>

What is wrong with the following "proof" of (∃x)F(x) ∴ (∀x)F(x)? Think about what the sequent would mean if the proof were valid.

1	(1)	(∃x)F(x)	A
2	(2)	F(a)	A
1	(3)	F(a)	1,(2,2) EE
1	(4)	(∀x)F(x)	3 UI

<u>Exercise: Predicate proofs</u>

1. Prove the equivalence of (∀x)F(x) and ~(∃x)~F(x). This requires you to construct a proof for (∀x)F(x) ∴ ~(∃x)~F(x) and a separate proof for ~(∃x)~F(x) ∴ (∀x)F(x). (Both can be done with RAA style arguments in less than 10 lines.)
2. Prove that (∀x)P(x) & (∀y)Q(y) ≡ (∀x)(P(x) & Q(x)).
3. Prove that (∃x)P(x) v (∃x)Q(x) ≡ (∃x)(P(x) v Q(x)).
4. Prove the sequent (∀x)(F(x)⇒P) ∴ (∃x)F(x)⇒P. This is a somewhat surprising sequent and what's more surprising is that the two wffs are actually equivalent! Can you think of an interpretation which makes this equivalence clear?
5. Using the predicates C(x), W(x) and D(x), convert the following statements into predicate wffs and construct an argument to prove the conclusion.

   "Computer programs are either well-written or difficult to maintain. Computer programs are not all well-written. Therefore, there are some computer programs which are difficult to maintain."

Sequent Introduction

Sequent introduction can be used in predicate arguments as it was in propositional arguments, but now the notion of a substitution instance is slightly more complicated.

In an earlier exercise we proved the sequent P⇒Q, R⇒~Q ∴ P⇒~R. This sequent makes a general claim which is valid for any interpretation of P, Q and R. (One interpretation might be P - husband, Q - married, R - bachelor.) We will use this sequent in the following example.

Suppose we want to prove the validity of the following Argument -

Some girls like all boys
No girl likes any male chauvinist
∴ No boy is a male chauvinist

As usual, we first define a dictionary -

G(x) means x is a girl
B(x) means that x is a boy
L(x,y) means that x likes y
M(x) means that x is a male chauvinist

With this dictionary the argument can be expressed as the sequent -

$$(\exists x)(G(x)\&(\forall y)(B(y)\Rightarrow L(x,y))),\ (\forall x)(G(x)\Rightarrow(\forall y)(M(y)\Rightarrow\sim L(x,y)))\ \therefore\ (\forall x)(B(x)\Rightarrow\sim M(x))$$

1	(1)	$(\exists x)(G(x)\&(\forall y)(B(y)\Rightarrow L(x,y)))$	A	First premise
2	(2)	$G(a)\&(\forall y)(B(y)\Rightarrow L(a,y))$	A	Typical disjunct. a is free but y is still bound
2	(3)	$G(a)$	2 &E	
2	(4)	$(\forall y)(B(y)\Rightarrow L(a,y))$	2 &E	
2	(5)	$B(b)\Rightarrow L(a,b)$	4 UE	Note that we use a different proper name
6	(6)	$(\forall x)(G(x)\Rightarrow(\forall y)(M(y)\Rightarrow\sim L(x,y)))$	A	Second premise
6	(7)	$G(a)\Rightarrow(\forall y)(M(y)\Rightarrow\sim L(a,y))$	6 UE	
2,6	(8)	$(\forall y)(M(y)\Rightarrow\sim L(a,y))$	7,3 MPP	
2,6	(9)	$M(b)\Rightarrow\sim L(a,b)$	8 UE	
2,6	(10)	$B(b)\Rightarrow\sim M(b)$	5,9 SI $P\Rightarrow Q,\ R\Rightarrow\sim Q\ \therefore\ P\Rightarrow\sim R$	
2,6	(11)	$(\forall x)(B(x)\Rightarrow\sim M(x))$	10 UI	Note that "b" does not occur in either assumption 2 or 6
1,6	(12)	$(\forall x)(B(x)\Rightarrow\sim M(x))$	1,(2,11) EE	Note that "a" does not appear in lines 11 or 6.

At line 10 we have used sequent introduction with substitution (see definition in the Propositional Logic section). The original sequent was $P\Rightarrow Q,\ R\Rightarrow\sim Q\ \therefore\ P\Rightarrow\sim R$, but we have substituted B(b) for P, L(a,b) for Q, and M(b) for R. Thus, "$B(b)\Rightarrow L(a,b),\ M(b)\Rightarrow\sim L(a,b)\ \therefore\ B(b)\Rightarrow\sim M(b)$" is a substitution instance of "$P\Rightarrow Q,\ R\Rightarrow\sim Q\ \therefore\ P\Rightarrow\sim R$" and whatever proof was used to establish the second could be modified to establish the first simply by making the appropriate substitutions. There is no trickery here, we are just saving time and ink.

> Exercise: Categorical Syllogisms
>
> A categorical syllogism is a three-line argument where each line is in one of the following forms -
>
> All X are Y
> Some X are Y
> No X is Y
> Some X are not Y
>
> 1. Express each of these four forms as predicate wffs.
> 2. Prove or disprove the following syllogisms -
>   i) All artists are extroverts
>   Some artists are poor people
>   ∴ Some poor people are extroverts
>   ii) All successful business managers work long hours
>   No one who works long hours has time for their family
>   ∴ No successful business manager has time for their family
>   iii) Some university courses are not computer science courses since some university courses are not fun whereas all computer science courses are fun.

Consistency and Completeness

In propositional logic we could test whether a wff was tautologous by examining its truth table or its PMD. This was the foundation on whose basis we established an equivalence between derivable sequents and tautological wffs. Thus we claimed that propositional logic is both complete and consistent. However, truth tables and PMDs do not work with wffs which contain predicates. If we want to determine whether predicate logic is complete and/or consistent then so we must re-examine our definition of tautology.

We want to show that -

i) If a predicate sequent can be proved using the 14 Rules of Derivation then the sequent is necessarily valid (Consistency); and

ii) For any valid deduction which can be expressed as a predicate sequent, there exists a formal argument which proves that sequent (Completeness).

Recall that truth tables and PMDs examine the truth of a wff in each possible model. In some parallel way we need to examine the validity of a predicate sequent under any possible model.

Interpretations

Any particular wff may be interpreted in different ways.

eg_1 (∀x)(F(x)v~F(x)) may mean "everything in the universe is either a football player or not a football player", or it may mean "everything in the universe is either green or not green".

eg_2 (∃x)(P(x)&~P(x)) may mean "there exists something which is both a pig and not a pig" etc.

Notice that for eg_1, any interpretation you invent will be true, but for eg_2, every interpretation will be false. This is the case regardless of what constitutes the universe: in any imaginable universe eg_1 will always be true and eg_2 false. (eg Even if you dream about a universe in which there are only computers and computer sales-people, and in which all sales-people wear red shoes, however you interpret the predicate F, (∀x)(F(x)v~F(x)) will turn out to be true.)

<u>Definition: Interpretations</u>

1. To create an <u>interpretation</u> of a wff (or sequent), one must define a universe of discourse and assign a meaning to each predicate and variable in the wff (or sequent).
2. If a <u>wff</u> is true under all interpretations in all non-empty universes, we say the wff is tautological (or necessarily true).
3. A <u>sequent</u> is valid if, under every interpretation in every non-empty universe, when the premises are true the conclusion is also true.

Consequently, if you can find any interpretation of a sequent which is not true, then the sequent is not valid. Note that to <u>prove</u> a sequent, you must construct a formal argument (since you cannot test each of the infinite number of possible interpretations) but to <u>disprove</u> a sequent, you just need to find one counter-example.

With this understanding of truth, validity and tautology, it can be shown that predicate logic is both consistent and complete.

<u>Exercise: Disproving sequents</u>

1. Find an interpretation which will show the following sequents to be invalid -
   - i) (∃x)F(x) ∴ (∀x)F(x)
   - ii) (∀x)(F(x)⇒G(x)), (∃x)G(x) ∴ (∃x)F(x)
   - iii) (∃x)(F(x)&G(x)), (∀x)(F(x)vG(x)) ∴ (∀x)G(x)
2. Prove or disprove the following sequents -
   - i) ∴ (∀y)(F(y)⇒(∃x)F(x))
   - ii) ∴ (∃y)(F(y)⇒(∀x)F(x))

Identity

Consider the simply deduction -

Only Smith and the guard knew the password.
Someone who knew the password stole the gun.
∴ Either Smith or the guard stole the gun.

This is clearly a valid argument, but try as you might you won't be able to proof it with the symbolism we have used so far: the quantities "all" and "some" are inadequate for handling this. We need some way of representing the claim that two things (say Smith and the person who stole the gun) are identical. So let us introduce a predicate I(x,y) to mean "x is identical to y". Such a predicate is immensely useful: in fact it is all we need to be able to express any natural number.

For instance, $(\exists x)H(x)$ may represent the claim "there is at least one human", but to represent "there are at least two humans" requires Identity. The wff $(\exists x)H(x)\&(\exists y)H(y)$ is inadequate since x and y may represent the same object. $(\exists x)(\exists y)(H(x)\&H(y))$ fails for the same reason. But using Identity we can write -

$$(\exists x)(\exists y)((H(x)\&H(y))\&\sim I(x,y))$$

("there is an x and a y which are both human and are not identical")

By common convention, we tend to write (x=y) instead of I(x,y) and hence "there is at least two" is written -

$$(\exists x)(\exists y)((H(x)\&H(y))\&\sim(x=y))$$

Similarly, "there is exactly one" could be written -

$$(\exists x)(H(x)\&\sim(\exists y)(H(y)\&\sim(x=y))$$

("there is at least one human, but there is no other human which is different than the first")

With this new notation, we can represent the stolen gun argument as -

$$(\forall x)(K(x)\Rightarrow((x=s)v(x=g)),\ (\exists x)(K(x)\&S(x))\ \therefore\ S(s)vS(g)$$

where K(x) means "x knows the password"
S(x) means "x stole the gun"
s stands for Smith
g stands for the guard

Having introduced this new notation it is then necessary to define two extra rules governing the introduction and elimination of identity (=I and =E). The proof of the stolen gun argument would then be straight-forward, but we will not bother with these details.

Exercise: Identity

Using Identity, express the following statements -

i) There are exactly two humans
ii) There are at least three different data types in Pascal
iii) There is more than one way to skin a cat
iv) There are two sides to every coin

Conclusion

Formalising logic as a system of propositions, logical connectives, predicates and variables may be viewed in two ways. First, it provides a systematic way of codifying human reasoning. Second, its formality is so rigidly defined that the whole system can be expressed and processed by computer. Thus, formal logic may be used as a bridge between human reasoning in natural language and automated reasoning by computer.

Computer programs have been written to construct truth tables and to check the validity of a proof. Unfortunately, the task of constructing proofs requires a certain degree of imagination and cannot be reduced to a mechanical procedure. Programs do exists which automate the proving of sequents, but all such programs are necessarily only partially successful. It is possible to keep generating longer and longer proofs in the hope of finding one which deduces the required sequent: first generate all one-line arguments, then all two-line arguments etc. Eventually a computer may be able to generate all arguments up to 100 lines in length and check whether any of them prove the desired sequent. Eventually the computer might be successful in finding a proof, but what happens if no such proof exists because the sequent is invalid? The computer will not be able to know whether the sequent is invalid or if there is still a proof which it hasn't yet found.

Section 7 - Other Forms of Logic

Propositional logic could not express all that we needed for logic, so we added predicates and quantifiers. Even then we could not express everything about quantities without introducing a special Identity predicate. It is still the case, however, that predicate notation is inadequate for expressing all of logic.

We turn now to four other approaches which each make some types of argument easier to express and validate.

Modal Logics

Modal logic is an attempt to capture the notions of necessity and possibility. Rather than say that every proposition must be either true or false, modal logic allows us to say "P must be true" and "Q may be true".

A sentence is *necessarily true* if it is true in all possible worlds and a sentence is *possible* if it is true in at least one possible world. The concept of a possible world is roughly the same as an interpretation in a non-empty universe. There are some imaginable universes, however, which are not possible worlds - for instance I can imagine a universe in which "I am alive" and "I am not alive" are simultaneously true (maybe in some nightmare!) but such a universe is not possible because it violates logic (in this case the Law of Non-Contradiction).

> Definition: Modal Operators
> i) $\Box$ W means that the wff W is necessarily true (it must be true, it is true in all possible worlds)
> ii) $\Diamond$ W means that the wff W is possibly true (it may be true, there is at least one possible world in which W is true)

These two operators are clearly inter-related -

i) $\Box W \equiv \sim\Diamond\sim W$

ii) $\Diamond W \equiv \sim\Box\sim W$

The following wffs are also fairly obvious -

iii) $\Box W \Rightarrow W$ (Axiom of Necessity - if something has to be true, then it is true)

iv) $W \Rightarrow \Diamond W$ (Axiom of Possibility - if something is true, then it is possible that it could be true)

v) $\Box W \Rightarrow \Diamond W$ (Combining iii and iv)

What about $\Diamond W \Rightarrow \Box W$? This claims that because something *could* be true, it *has to* be true. But this is clearly invalid - for instance my name is Matthew and hence it is possible that my name is Matthew (by iv), but it doesn't follow that my name has to be Matthew (since in another possible world I may be called some other name).

Examples of Modal Logic

There is not just one "modal logic", but a variety of modal logics which each interpret the two operators $\Diamond$ and $\Box$ differently.

1. Temporal Logic

Definition: Temporal logic

$\Box$ is interpreted to mean "always true at every time in the future"

$\Diamond$ is interpreted to mean "may be true at some time in the future"

eg_1 "If you are dead than you will always remain dead" - if we let the symbol D represent being dead, then this sentence could be written as the modal wff -

$$D \Rightarrow \Box D$$

eg_2 The opposite of eg_1 might be "It is conceivable that someone might come back from the dead", which could be written as -

$$D \& \Diamond D$$

eg_3 "There may be a rainy day in the future." In the universe of all days, where R(x) means "x is a day in which it rains", we can write -

$$\Diamond (\exists x)R(x)$$

2. Physical Logic

Definition: Physical logic

$\Box$ is interpreted to mean "in the real universe the laws of physics require that ..."

$\Diamond$ is interpreted to mean "in the real universe it could happen that ..."

eg_1 "It is not possible to win a lottery unless you buy a ticket." Let B represent buying a lottery ticket and W represent winning the lottery. Then we can write -

$$\sim B \Rightarrow \sim \Diamond W$$

As with our previous systems, one statement may be represented equally well by several different wffs. For instance the last example could easily be written $\sim B \Rightarrow \Box \sim W$

eg_2 "It is impossible to travel faster than light." Let S(x,y) mean that object x is travelling at speed y; G(x,y) mean that x>y; and c stand for the speed of light. Then we can write -

$$\sim \Diamond (\exists x)(\exists y)(S(x,y) \& G(y,c))$$

3. Epistemic Logic

Definition: Epistemic logic - ie the logic related to knowledge

$\Box$ is interpreted to mean "it is commonly known that..."

$\boxed{t}$ is interpreted to mean "The person t knows that..."

$\Diamond$ is interpreted to mean "it is commonly thought that ... may be true"

$\Diamond_t$ is interpreted to mean "The person t thinks that ... may be true"

eg_1 "Everyone knows that Themba is a boy." Let B(x) mean that x is a boy; and let t stand for Themba. Then this sentence could be written as the modal wff -

$$\Box B(t)$$

eg_2 "Sue thought that maybe Themba didn't know she was rich." Let R(x) mean that x is rich; let s stand for Sue; and let t stand for Themba. Then this sentence could be written as the modal wff -

$$\Diamond_s \sim \boxed{t} R(s)$$

4. Other forms

A number of other forms of modal logic exist such as Moral Logic ("it is morally imperative that...") and Action Logic ("after a particular action it will certainly be true that ...").

> Exercises: Modal Logics
>
> For each of the following sentences, choose which type of modal logic is most applicable and represent the sentence as a modal wff.
>
> i) There will never be a day when everybody agrees completely, but there is still hope that one day we will live in peace.
>
> ii) Themba did not realise that his wife was ugly.
>
> iii) The boiling point of any substance must be higher than its freezing point.

Arguments with modal wffs

Various systems have been proposed to take wffs like those above and include them in formal arguments.

1. System T

> Definition: System T
>
> Take all of predicate logic and add -
>
> - $\Diamond W \equiv \sim\sim W$
> - A $\therefore$ B is a valid sequent iff $(A \Rightarrow B)$ is a tautology
> - $P \Rightarrow P$ is a tautology
> - $(P \Rightarrow Q) \Rightarrow (P \Rightarrow Q)$ is a tautology
> - If P is a tautology then so is P

(Note that the last addition is <u>not</u> the same as saying that $P \Rightarrow P$ is a tautology. If P is <u>always</u> true then P is true; but just because P <u>happens</u> to be true does not mean that P is true.)

2. System S_4

Take all of System T and add -

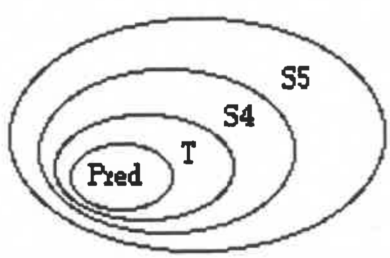
- $P \Rightarrow P$ is a tautology

3. System S_5

Take all of System T and add -

- $\Diamond P \Rightarrow \Diamond P$ is a tautology

Both S_4 and S_5 are based on T, but $P \Rightarrow P$ can be derived from $\Diamond P \Rightarrow \Diamond P$ and hence S_5 is a more powerful system than S_4.



Multi-Valued Logics

One of the first assumptions of propositional and predicate logic is that every wff must be either true or false - there is no middle ground. (This is stated explicitly by the Law of Excluded Middle.) But is this always the case? Consider the following sentences -

i) My mother's phone number is 971 333
ii) The population of South Africa right now is exactly 37,659,051
iii) Professor Finnie is tall

In the first example, the reader probably doesn't know whether the claim is true or false, but we could still say that regardless of whether we know the truth of a sentence, it still must be either true or false. However, nobody will ever know whether the second sentence is true or false - it is physically impossible to ever find out. Do we still want to claim that a sentence must be either true or false when it is impossible in principle to ever be able to tell which?

The third example is more problematic since the predicate "tall" is vague. Even if we knew Prof. Finnie's exact height there will be disagreement about whether the claim is true or false.

Three-Valued Logic

Instead of restricting truth values to T and F, a three-valued logic also has a third, intermediate value. This third value is represented by I, which may mean "unknown", or "possibly true" etc.

To build a solid basis for logic with three values, we first need to show how the logical connectives ~,&,v,⇒ are defined for those values. That is, the truth tables for two-valued logic will have to be modified. While constructing these tables, it is useful to think of I as "not yet known", however, we will see that this leads to some trouble.

A	~A
T	F
I	I
F	T

A B	A & B
T T	T
T I	I
T F	F
I T	I
I I	I
I F	F
F T	F
F I	F
F F	F

> Exercises: Three-valued truth tables
> Construct similar truth tables for AvB and A⇒B.

Using truth table definitions based on the "not yet known" interpretation of I leads to the following problems. The wff Pv~P has always been treated as tautologous: it is true regardless of the truth of P. We would hope that even if the truth of A was I, then the wff overall should be true. But the truth table definition states that if A is I and B is also I then AvB is I as well. That is, there is a condition under which Pv~P is not true.

Similarly, P⇒P has always been tautologous before, but according to the "not yet known" truth table, when P is I the overall wff is also I.

These anomalies arise because we aren't clear about exactly what the intermediate value I means. It would appear that "not yet known" is not a good choice. We could alter the definitions to correct such anomalies, but any change would bring new problems. Nevertheless it is common to make one adjustment - in the table for implication, if A is I and B is I then we define A⇒B to be T. This alteration simplifies the mathematical analysis in the next section, but has the unfortunate side effect that A⇒B is no longer equivalent to ~AvB.

More than three values

The concept of creating an intermediate value can be generalised to allow any number of truth values between T and F.

> Definition: Multi-valued logic
> Suppose we allow M different truth values numbered from 1 (absolutely false) to M (absolutely true). These numbers represent varying degrees of assertability.
>
> Pick some number S between 1 and M so that truth values from 1 to S mean "false to some extent" and those from S+1 to M mean "true to some extent". For any propositional wff we want to be able to calculate its truth value and if the value is ≤S we will deny the wff, but if it is >S we will assert the wff.

eg With M=5 and S=2 the truth values might be -
1 - Definitely not
2 - Maybe not
3 - Most likely
4 - Almost certainly
5 - Definitely

What would a truth table look like for negation in a five-valued logic?

A	~A
1	5
2	4
3	3
4	2
5	1

In general, if V(P) is the truth value of some wff P in a M-valued logic, then negation is defined by the mathematical expression -

$$V(\sim P) = M-V(P)+1$$

The truth of P&Q will be the same as the truth of its most doubtful conjunct. Hence -

$$V(P\&Q) = \min(V(P), V(Q))$$

Similarly we can write mathematical expressions for the other connectives -

$$V(PvQ) = \max(V(P), V(Q))$$
$$V(P\Rightarrow Q) = \min(M, V(Q)-V(P)+M)$$

> Exercises: Two-valued truth tables
> Study the mathematical expressions for calculating truth values in multi-valued logics. Show that when M=2, these calculations give exactly the same answers as standard truth tables.

Probabilistic Logics

The probability of some hypothesis H being true can be expressed as a real number between 0 and 1. This is actually just an extension of the concept of multi-valued logic. Instead of having a finite number of truth values from 1 to M, we have an infinite number of truth values between 0 and 1.

In many computing applications (especially expert systems which try to cope with uncertain data) such probabilities are used to express the degree to which some proposition is believed to be true.

Definition: Probabilistic connectives	
Negation	P(~H) = 1-P(H)
Conjunction	P(A&B) = P(A).P(B)
Disjunction	P(AvB) = P(A)+P(B)-P(A).P(B)

The expression for disjunction may be confusing, but think of it in set terms. Suppose A and B are overlapping sets and we are counting the number of elements in their union. Then we will add the number in A and the number in B, but this will have counted those in the intersection twice, so we then subtract the number in the intersection of A and B.

eg If the proposition "Themba was awake at 3am this morning" is believed with certainty 0.85, we may write P(A(t)) = 0.85 (where A(x) is the predicate "was awake at 3am"). Then we could decide how much to believe the proposition "Everybody was asleep at 3am this morning"?
First we write the proposition as the wff $(\forall x)\sim A(x)$ which we note is in direct contradiction to the other proposition A(t): the two claims are mutually exclusive. Hence $P((\forall x)\sim A(x)) = P(\sim A(t)) = 1-P(A(t)) = 0.15$

Exercises: Probabilistic connectives

Suppose "Themba is married" (M) is believed to the degree of 0.5 and that he has "never loved an ugly girl" (U) to the degree of 0.9.

i) What is our degree of believe in (M&U)?
ii) What is our degree of believe in (MvU)?

Bayes Theorem

One commonly used piece of probability theory is a theorem which helps us to calculate the probability of some hypothesis H given some knowledge about related events.

Definition: Bayes Theorem

$$P(H \mid e) = \frac{P(e \mid H).P(H)}{P(e)}$$

This equation can be seen as updating the belief in H given some new evidence e. Suppose that H is "my house has just been burgled" and that police statistics show that one in every 100 houses are burgled in your area every night. ie P(H)=0.01. If your friend Themba rings and says that your house is being burgled, should you believe him? How will our belief in H change?

Suppose P(Themba rings | burglary actually occurred) = 0.9 but that P(Themba rings | no burglary has occurred) = 0.35 (since Themba likes playing practical jokes). Suppose also that we calculate (or reason) from other sources that P(Themba rings) = 0.4.

Then P(Burglary occurred | Themba rings) = (0.9*0.01) / 0.4 = 0.0225. ie this new evidence increases your belief in H from 0.01 to 0.0225.

Certainty Factors

Definition: Certainty factors

A certainty factor is a real number between -1 and 1: negative numbers indicate degree of disbelief in some proposition; 0 means no evidence is available for or against the proposition; positive numbers indicate degree of belief.

-1	0	+1
Total disbelief	No evidence	Total belief

eg Themba phones and says that the burglar alarm on your house has gone off. Not quite knowing whether to trust Themba, you ring another neighbour Mrs Govender and she confirms that she can hear your alarm.

You reason as follows -

1. If Themba says the alarm went off, you would believe it with certainty $CF_1=0.5$.
2. If Mrs Govender says she hears something, you would believe it with certainty $CF_2=0.9$.
3. Given both 1 and 2, you then combine the two CFs to calculate a CF for your belief that the alarm sounded -

$$CF_3=CF_1+CF_2 - CF_1.CF_2 = 0.95$$

Definition: Parallel Combination Function

When two rules support the same conclusion, the certainty factors are combined using the formula -

$$CF_3 = \begin{cases} CF_1+CF_2-CF_1.CF_2 & \text{if } CF_1 \text{ and } CF_2 \geq 0 \\ CF_1+CF_2+CF_1.CF_2 & \text{if } CF_1 \text{ and } CF_2 \leq 0 \\ \dfrac{CF_1+CF_2}{1-\min(|CF_1|,|CF_2|)} & \text{otherwise} \end{cases}$$

4. If the alarm sounds then you believe there is a burglary with certainty $CF_4=0.99$. Given that you believe the alarm *has* sounded with certainty $CF_3=0.95$, to what extent do you believe there has been a burglary?

$$CF_5=CF_3.CF_4 = 0.94$$

Definition: Serial Combination Function

When the hypothesis supported by one rule becomes evidence for a second rule, the certainty factors are combined using the formula -

$$CF_5 = \begin{cases} CF_3.CF_4 & \text{if } CF_3 > 0 \\ 0 & CF_3 \leq 0 \end{cases}$$

It is also possible to handle conjunctions and disjunctions in a system which uses certainty factors. This would involve calculations using min and max similar to those described in the section above on multi-valued logics. All of these formulae for combining certainty factors can be expressed in terms of Bayes Theorem.

Fuzzy Logic

Fuzzy logic is another form of multi-valued logic designed to capture vague terms such as large profit, high pressure, tall, very, a few, most etc. It would be possible to give strict definitions of such terms (eg define "tall" to mean "at least 180cm"), but this seems artificial and loses the flexibility inherent in everyday conversation. We want such terms to be fuzzy rather than exact.

Fuzzy Sets

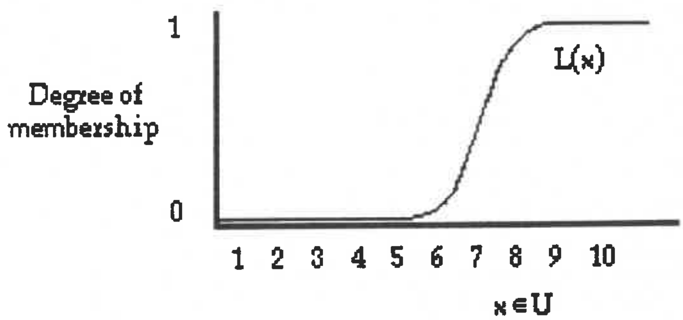
In set theory an object is either a member of a set or it is not. For fuzzy sets this boolean characterisation is replaced with a measure of the degree of membership ranging from 0 (not a member) to 1 (completely or definitely a member).

Let the universe be the set U and let A be a fuzzy subset of U. Define a membership function A(x) which gives a value between 0 and 1 for each $x \in U$.

eg U = {1,2,3,...10} and L = large numbers
Member of L is not simply true or false, but rather is defined by the function -

$$L(x) = \begin{cases} 0 & \text{if } x \leq 5 \\ 0.2 & \text{if } x=6 \\ 0.5 & \text{if } x=7 \\ 0.8 & \text{if } x=8 \\ 1 & \text{if } x \geq 9 \end{cases}$$

Fuzzy set membership is often shown as a graph, which in this case looks like -



Various facts

i) A fuzzy set is empty ($A=\emptyset$) iff $A(x)=0$ for all $x \in U$.
ii) $A \equiv B$ iff $A(x)=B(x)$ for all $x \in U$.
iii) $A \subseteq B$ iff $A(x) \leq B(x)$ for all $x \in U$.
iv) $|A| = \sum_{x \in U} A(x)$
v) If $A(x)$ is the membership function for a fuzzy set A then the membership function for ~A is $1-A(x)$.
vi) The membership function for $A \cap B$ is $\min(A(x), B(x))$.
vii) The membership function for $A \cup B$ is $\max(A(x), B(x))$.

Converting a fuzzy set into a normal set
Any fuzzy set can be compressed into a normal set by declaring a threshold. For instance, a computer program may manipulate a variety of fuzzy sets, but often the user does not want output in the form of a fuzzy value. That is they do not want to see "8 is a large number" expressed as a degree of membership of (0.8) as output, they simply want to see YES or NO.

Define the set Large = $\{x \mid L(x) \geq 0.75\}$. By setting the number 0.75 as a threshold, it now becomes true that $8 \in$ Large.

Definition: α-Cut
In general an α-cut is a (normal) set whose elements are all those objects whose degree of membership of the fuzzy set is at least as great as the threshold α.

$$A_\alpha = \{ x \mid x \in U \text{ and } A(x) \geq \alpha \}$$

Fuzzy Logical Connectives

Can we define the traditional logical connectives as fuzzy operations?

$$\text{Suppose Tall}(x) = \begin{cases} 0 & \text{if } x \leq 150 \\ \frac{x-150}{30} & \text{if } 150 < x \leq 180 \\ 1 & \text{if } x > 180 \end{cases}$$

What is the truth value of a proposition like "Rob is tall"?
V("Rob is tall") = Tall(Rob's height) = Tall(175) = 0.71
V("Bart Simpson is tall") = Tall(Bart's height) = Tall(120) = 0

Negation: V("Rob is not tall") = V(~"Rob is tall") = 1-V("Rob is tall") = 0.29

Conjunction: V("Rob and Bart are both tall") = min(V("Rob is tall"), V("Bart is tall")) = 0

Disjunction: V(Either Rob or Bart is tall") = max(V("Rob is tall"), V("Bart is tall")) = 0.71

Implication: V("If Rob is tall then Bart is tall") = min(1, V("Bart is tall")-V("Rob is tall")+1)

= min(1, 0.29) = 0.29

Fuzzy Quantification

Not only are predicates (such as "tall") fuzzy, but also quantifiers (such as "most", "many", "a few" etc) can be fuzzy.

eg Express the fuzzy statement "Most of my friends are honest."

Let Friends = {Themba, Sue, Fred}.

Define a fuzzy predicate for honesty, and suppose that honest(Themba)=0.1, honest(Sue)=0.6 and honest(Fred)=0.8.

A fuzzy function can be defined for any quantifier if we can first calculate the ratio -

$$r = \frac{\text{the degree to which objects satisfy some criterion}}{\text{the total number of objects}}$$

In our example we could calculate the expected honesty of my friends as the average -

$$\frac{\sum_{x \in \text{Friends}} \text{honest}(x)}{|\text{Friends}|} = \frac{(0.1+0.6+0.8)}{3} = 0.5$$

Given such a ratio, we could define define the quantifier "all" very strictly -

$$\text{All}(r) = \begin{cases} 0 & \text{if } r<1 \\ 1 & \text{if } r=1 \end{cases}$$

But the quantifier "most" can be defined more loosely -

$$\text{Most}(r) = \begin{cases} 0 & \text{if } r\leq 0.3 \\ 2r-0.6 & \text{if } 0.3<r<0.8 \\ 1 & \text{if } r\geq 0.8 \end{cases}$$

With these definitions, V("Most of my friends are honest") = Most(0.5) = 0.4

Bibliography

Barker, S.F.; The Elements of Logic; McGraw-Hill, 1965

Bradley, R. and Swartz, N.; Possible Worlds: an introduction to logic and its philosophy; Blackwell, 1979

Bucker, R.; Infinity and the Mind; Birkhäusser, 1982

Copi, I.M.; Introduction to Logic; MacMillan, 1982

Fearnside, W.W.; About Thinking; Prentice-Hall, 1980

Hofstadter, D.R.; Gödel, Escher, bach: an eternal golden braid; Penguin, 1980

Hughes, G.E. and Cresswell, M.J.; An Introduction to Modal Logic; Methuen 1968

Lemmon, E.J.; Beginning Logic; Nelson's University Paperbacks, 1971.

McCawley, J.D.; Everything that linguists have always wanted to know about logic but were ashamed to ask; Blackwell, 1981