

EDWIN D. MARES

Relevant Logic

A Philosophical Interpretation

CAMBRIDGE

Relevant Logic

This book introduces the reader to relevant logic and provides the subject with a philosophical interpretation. The defining feature of relevant logic is that it forces the premises of an argument to be really used ('relevant') in deriving its conclusion. The logic is placed in the context of possible world semantics and situation semantics, which are then applied to provide an understanding of the various logical particles (especially implication and negation) and natural language conditionals. The book ends by examining various applications of relevant logic and presenting some interesting open problems. It will be of interest to a range of readers including advanced students of logic, philosophical and mathematical logicians, and computer scientists.

EDWIN D. MARES is a Senior Lecturer in the Philosophy Programme, Victoria University of Wellington. He has published extensively on both the philosophical and mathematical aspects of logic, as well as on metaphysics and the philosophy of language.

Relevant Logic

A Philosophical Interpretation

Edwin D. Mares



CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press

The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521829236

© Edwin D. Mares 2004

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2004

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Mares, Edwin David.

Relevant logic: a philosophical interpretation / Edwin D. Mares.

p. cm.

Includes bibliographical references and index.

ISBN 0 521 82923 2 (hardback)

1. Relevance logic. I. Title.

BC41.M27 2004

160 – dc22 2003047253

ISBN-13 978-0-521-82923-6 hardback

ISBN-10 0-521-82923-2 hardback

Transferred to digital printing 2007

Contents

<i>Preface</i>	vii
<i>Acknowledgements</i>	x
I Relevant logic and its semantics	1
1 What is relevant logic and why do we need it?	3
2 Possible worlds and beyond	19
3 Situating implication	39
4 Ontological interlude	57
5 Negation	73
6 Modality, entailment and quantification	96
II Conditionals	123
7 Indicative conditionals	125
8 Counterfactuals	144
III Inference and its applications	161
9 The structure of deduction	163
10 Disjunctive syllogism	176
11 Putting relevant logic to work	189
12 Afterword	207

Appendix A	The logic R	208
Appendix B	Routley–Meyer semantics for R	210
<i>Glossary</i>		216
<i>References</i>		218
<i>Index</i>		226

Preface

This book is a philosophical interpretation of relevant logic. Relevant logic, also called ‘relevance logic’, has been around for at least half a century. It has been extensively developed and studied in terms of its mathematical properties. So relevant logic is a highly developed and mathematically well-understood branch of non-classical logic. But what is it good for and why should we adopt it? I think that it is a good tool for understanding ordinary deductive reasoning and that it provides us with the tools to understand conditionals. And that is what this book is all about.

Unlike intuitionist logic, relevant logic does not come packaged with its own philosophy. There are intuitionists, and they all share a large number of important philosophical views that non-intuitionists reject. Although some relevant logicians have talked about ‘the relevantist’, relevantism is not a well-developed view, nor one that is widely held even by relevant logicians. By and large, we are free to adopt their own philosophical interpretation of relevant logic.

Historically, my own view developed out of my acquaintance with the possible worlds approach to semantics. When I was a graduate student, I studied modal logic and Montague grammar and found the framework of possible worlds to be a very intuitive, elegant and powerful framework in which to do semantics. Later, after I had become immersed in relevant logic, I wanted to give others the same sort of feeling of being at home in relevant logic that I had felt when I was first exposed to possible world semantics. This book is the latest product of that attempt.

I begin with the possible worlds framework. It provides both the basis for my semantics and the ontology that underpins that semantics. The central notion in my interpretation of relevant logic is that it shows us how to make inferences about what *situations* hold at worlds. I have borrowed the notion of a situation from the situation semanticists – Jon Barwise, John Perry and their followers. But I have constructed these situations out of the elements available in possible world semantics.

What I want to do in this book is be able to give philosophers and others an intuitive grasp of what’s going on in relevant logic and its semantics, and to show that it is viable and useful.

I have three main aims in this book. First, I try to show that relevant logic and its semantics are intelligible and intuitive. Second, I argue that relevant logic is useful. It provides us with a theory of inference, is the basis for a good theory of conditionals, and has several other uses. Third, I try to demonstrate that relevant logic does not force any untoward philosophical commitments on us. I argue this last point by constructing a model for relevant logic out of possible worlds, individuals and structures available in Peter Aczel's non-well-founded set theory. The latter is a very elegant theory and I recommend it to everyone working in logic.

Possible world semantics is the current paradigm in philosophical logic. If I can show that the same entities (pretty much) that are used in possible world semantics can be used to construct the model that I intend for relevant logic, I can undercut any metaphysical argument against relevant logic.

Although I do not assume any previous familiarity with relevant logic, this book is not a textbook in relevant logic. Rather, it is a philosophy book about relevant logic. I have tried to make the ideas in it accessible to people familiar with natural deduction. I have not tried to provide a survey of relevant logics, nor of the mathematical results about relevant logic. There are various good surveys of this sort. Among the shorter surveys are (Dunn 1986), (Mares and Meyer 2001), and (Mares 2002b). For readers with a good deal of logical sophistication, there are the *Entailment* volumes (Anderson and Belnap 1975) and (Anderson *et al.* 1992), and the volumes of *Relevant Logics and their Rivals* (Routley *et al.* 1982) and (Brady 2003). More recently, there are two excellent textbooks on substructural logics. Relevant logics are substructural logics, and so are treated in these books. The books are Greg Restall's *An Introduction to Substructural Logics* (Restall 2000) and Francesco Paoli's *Substructural Logic: A Primer* (Paoli 2002). I recommend Restall's and Paoli's books for readers who want to explore the technical properties of relevant logics and compare them to other sorts of logics. For the reader who would like a good and very readable general introduction to the topic of non-classical logics, there is Graham Priest's *An Introduction to Non-Classical Logic* (Priest 2001).

One last note. I use the term 'relevant logic' rather than 'relevance logic' because it sounds more natural. I find that North Americans tend to use 'relevance logic' and Australasians and Europeans use 'relevant logic'. I haven't canvassed the views of people from other continents, so I can't speak with authority about what they would say. Some people have tried to link different attitudes towards logic to the use of these different terms, but my usage is merely one of convenience. In fact, in an early draft of this book I used the two terms interchangeably. I was told by my editor and three referees to choose one term and stick with it. So I chose 'relevant logic'.

The plan of this book

The book is divided into three parts. In the first, I outline relevant logic and its model-theoretic semantics. While doing so, I give both philosophical motivations. In so doing, I do not discuss every aspect of the logic and its semantics. Rather, I paint a picture of the logic using a fairly broad brush. I leave ‘house-keeping’ details for appendices that appear at the end of the book. I have a chapter on each of the more problematic connectives – implication, negation, and modal operators. In addition, there is a chapter introducing the principal ideas behind relevant logic, a chapter introducing possible world semantics and the Routley–Meyer semantics, and a chapter discussing the metaphysical implications of adopting relevance logic.

The second part uses the semantics and ideas from the first part to develop a theory of conditionals. There are two chapters in this section – one on indicative conditionals and the other on counterfactuals.

The third part of the book goes into more detail about the theory of deductive inference. Chapter 9 gives some technical details concerning the nature of premises and logical consequence. Chapter 10 uses these details, along with the theory of conditionals developed in part II, to give an analysis of the rule of disjunctive syllogism. In the final chapter, further uses of relevant logic and some interesting open problems are discussed.

Acknowledgements

I have discussed issues that crop up in this book with almost everyone I know and most of the people I have met. I can't list all of their names, but I will try to get in everyone with whom I have had substantive discussions or who has read drafts of this book. Among these, three stand out. Mike Dunn first taught me relevant logic when I was a graduate student. I have learned a great deal from his own work and from his comments on my work. Bob Meyer was my mentor when I was a Visiting Fellow at the Automated Reasoning Project at the Australian National University. We wrote a series of papers together and continue to work together on projects in relevant logic. I cannot express how much I have learned from Bob or the debt of gratitude that I feel I owe him. Max Cresswell read an early draft of this book and made very valuable comments. He also talked about logic (relevant, modal, and otherwise) with me on a great many occasions and I have also learned much from him.

Others who have helped me to write this book are Nick Agar, John Barwise, Kata Bimbó, Ross Brady, James Chase, Nino Cocchiarella, Kit Fine, André Fuhrmann, Lou Goble, Rob Goldblatt, Julianne Jackson, Neil Leslie, Bernie Linsky, Errol Martin, Charles Morgan, Adriano Palma, Jeff Pelletier, John Perry, Jane Pilkington, Graham Priest, Stephen Read, Greg Restall, Peter Schotch, Oliver Schulte, Jerry Seligman, John Slaney, Nick Smith, Koji Tanaka, Alasdair Urquhart, Ed Zalta, Neil de Cort and Susan Beer. Hilary Gaskin of Cambridge University Press has shown great care and patience in helping this project along from its beginning as a very messy and error-riddled manuscript.

On a more personal level I should thank my parents, Joseph and Martha Mares, and the rest of the Mares clan: Eric, Diana, Naomi, and Joel, who have encouraged me for years to finish this book. And I am grateful to my dog, Ramsey, who slept behind my chair, went for walks with me, and didn't object when I rattled on about logic. This book is dedicated to Ramsey's memory.

Part I

Relevant logic and its semantics

1 What is relevant logic and why do we need it?

1.1 Non-sequiturs are bad

The central aim of relevant logicians has been to give a more intuitive characterisation of deductive inference. Consider the following example. Since 1993, when Andrew Wiles completed his difficult proof of Fermat's Last Theorem, mathematicians have wanted a shorter, easier proof. Suppose when someone addressing a conference of number theorists, suggests the following proof of the theorem:

The sky is blue.

∴ There is no integer n greater than or equal to 3 such that
for any non-zero integers x, y, z , $x^n = y^n + z^n$.

This proof would not be well received. It is a non-sequitur – its conclusion does not, in any intuitive sense, follow from its premise. It is a bad proof.

But let's think about this a little harder. According to the standard definition of 'validity', an argument is valid if and only if it is impossible for the premises all to be true and the conclusion false. The conclusion is not just true; it is necessarily true. All truths of mathematics are necessary truths. Thus, it is impossible for the conclusion to be false. So, it is impossible for the premise to be true and the conclusion to be false. Therefore the argument is valid. Moreover, the proof is sound, since its premise is also true. But it is clear that it is a bad argument and no one would take it seriously. Therefore, we need some notion other than the standard definition of validity to determine what is wrong with this proof. I suggest that what is wrong is that the standard notion of validity is too weak to provide a vertebrate distinction between good and bad arguments. It allows too many non-sequiturs to be classified as good arguments.

The standard notion of validity is at the heart of the logical theory known as *classical logic*. Classical logic is the sort of logic that students learn in introductory symbolic logic courses. Since we deal almost exclusively with propositional logic in this book, it will suffice here to discuss only the classical propositional calculus. The method of determining validity that we teach first to students is the method of truth tables. We first list the propositional variables

that occur in the premises and conclusion and then we list each of the premises and the conclusion, all along a top row. Then we list in each subsequent row one combination of truth values for the propositional variables. The following is a schematic truth table for an inference with two propositional variables, premises $A_1, \dots, A_n$ and conclusion B :

p	q	$A_1, \dots, A_n$	B
T	T		
T	F	$\vdots$	$\vdots$
F	T		
F	F		

Each row of the truth table defines a ‘possibility’¹ and determines the truth value of each premise and conclusion in each of these possibilities. An argument is valid, according to the classical propositional calculus, if and only if in every row in which every premise is true the conclusion is also true.

Now, if we have a conclusion that is true in every possibility, then the argument is valid, regardless of what the premises say. Consider, then, a statement that we can prove to be true in every possibility on the truth tables, such as, $q \vee \sim q$. On the classical account of validity, the following argument is valid:

$$\frac{p}{\therefore q \vee \sim q}$$

Thus, if we set p to mean ‘my dog barks at rubbish collectors’ and q to mean ‘it is raining in Bolivia right now’ we find out that the inference

$$\frac{\text{My dog barks at rubbish collectors}}{\therefore \text{Either it is raining in Bolivia right now or it is not.}}$$

is valid. Since the premise is true, the argument is also sound. But like the first inference that we investigated, it is a very bad argument. Thus, the classical notion of validity does not agree with our pre-logical intuitions about where the division between good arguments and non-sequiturs should be. Classical logic allows connections between premises and conclusions in valid arguments that are extremely loose. There needs to be more of a connection between the content of the premises and conclusion in an argument that we are prepared to call ‘valid’.

¹ A possibility here is not a possible world. The rows of truth table are far too under-specified to determine a possible world. We will look at possible worlds in some depth in chapter 2 below.

1.2 The classical theory of consequence

We can avoid the problem given above if we think about the issue in syntactic, rather than semantic terms. In a mathematical proof, for example, we could demand that only obvious axioms and a few primitive rules are to be used. If we take this line, then the ‘proof’ of Fermat’s last theorem given above would be rejected. Similarly, we could demand that an argument such as the one from ‘My dog barks at rubbish collectors’ to ‘Either it is raining in Bolivia right now or it is not’ be recast in terms of a system of proof, such as a natural deduction system, that also contains only a few rules.

But there are still problems with the classical view of consequence, even treated in this form. To see these problems, let’s consider a standard proof theory for classical logic, of the sort that we teach to students in an introductory course on symbolic logic. The system that we consider here is a ‘Fitch-style’ natural deduction system (Fitch 1952). Readers who are familiar with this type of natural deduction system can skim this section. But don’t skip it altogether, for there are some points made in it that we will build on later.

The structures that we construct in a natural deduction system are proofs. The first step in a proof is a formula that is a hypothesis. Every subsequent step in the proof is also a hypothesis or it is a formula that is derived from previous steps by means of one of the rules of the system. Each hypothesis introduces a subproof of the proof. The notion of a proof and a subproof are relative. Subproofs themselves can have subproofs. Subproofs here are indicated by means of vertical lines, as we shall see below.

There are three rules of proof that we will use that do not involve any logical connective, such as conjunction, disjunction, implication, or negation. The first of these is a rule that allows us to introduce a formula as a premise, or hypothesis. The other two rules allow us to copy formulae in proofs. The rule of repetition (*rep*) allows us to copy a line of a proof within that proof and the rule of reiteration (*reit*) allows us to copy a line from a proof into any of its subproofs (or subproofs of subproofs, and so on). This will all become clearer when we look at some examples of proofs. But before we do so, we need to discuss the rules governing the connectives.

Each connective has two rules associated with it – an introduction rule and an elimination rule. In this chapter, we will only be concerned with the implication connective ($\rightarrow$). The elimination rule for implication is²

($\rightarrow E$) From $A \rightarrow B$ and A to infer B .

The rule of implication introduction is a little more difficult. It says that

($\rightarrow I$) Given a proof of B from the hypothesis A ,
one may infer $A \rightarrow B$.

² I present rules of inference in the style of (Anderson *et al.* 1992).

Consider, for example, a proof of the formula $A \rightarrow ((A \rightarrow B) \rightarrow B)$:

1.	A	hyp
2.	A $\rightarrow$ B	hyp
3.	A	1, <i>reit</i>
4.	B	2, 3, $\rightarrow$ E
5.	(A $\rightarrow$ B) $\rightarrow$ B	2 – 4, $\rightarrow$ I
6.	A $\rightarrow$ ((A $\rightarrow$ B) $\rightarrow$ B)	1 – 5, $\rightarrow$ I

Now let's take a look at an easier proof, this time for the scheme, $A \rightarrow A$:

1.	A	hyp
2.	A	1, <i>rep</i>
3.	A $\rightarrow$ A	1 – 2, $\rightarrow$ I

There seems to be nothing objectionable about this proof. But, according to the classical theory of consequence, we can convert it into a proof of $A \rightarrow A$ from the *irrelevant* hypothesis B as follows:

1.	B	hyp
2.	A	hyp
3.	A	2, <i>rep</i>
4.	A $\rightarrow$ A	2 – 3, $\rightarrow$ I

Thus, for example, we can prove from 'The sky is blue' that 'Ramsey is a dog implies that Ramsey is a dog.' The former has nothing to do with the latter.

What the above proof illustrates is that in the classical theory of consequence we can add any premise to a valid deduction and obtain another valid deduction. Thus we can add completely irrelevant premises. This seems to be far too generous a notion of consequence and the classical notion of implication seems to capture a far looser relationship between propositions than does our common notion of implication.

Let us go on to see how relevant logic is supposed to remedy this situation.

1.3 The real use of premises

The problem with non-sequiturs like the ones discussed above is that some of the premises of the inferences appear to have nothing to do with the conclusion. Relevant logic attempts to repair this problem by placing a constraint on proofs that the *premises really be used in the derivation of the conclusion*.

We will present this idea in the context of Anderson and Belnap's natural deduction system for relevant logic. The idea is pretty simple. Each premise, or rather hypothesis, in a proof is indexed by a number. The various steps in a proof are indexed by the numbers of the hypotheses which are used to derive

the steps. For example, the following is a valid argument in this system:

- $$\begin{array}{ll}
 1. & A \rightarrow B_{\{1\}} \quad \text{hyp} \\
 2. & A_{\{2\}} \quad \text{hyp} \\
 3. & A \rightarrow B_{\{1\}} \quad 1, \text{reit} \\
 4. & B_{\{1,2\}} \quad 2, 3, \rightarrow E
 \end{array}$$

The numbers in curly brackets are indices. The indexing system allows us to do away with the scope lines, since the numbers indicate in which subproof a step is contained. Removing the scope lines, we get the following proof:³

- $$\begin{array}{ll}
 1. & A \rightarrow B_{\{1\}} \quad \text{hyp} \\
 2. & A_{\{2\}} \quad \text{hyp} \\
 3. & B_{\{1,2\}} \quad 1, 2, \rightarrow E
 \end{array}$$

Here we can see, without scope lines, that there are two subproofs, since there are two indices (hence two hypotheses) used in the proof. Moreover, we do not need to use the reiteration rule in this proof. The indices make sure that when we use a step from a proof and one from a subproof in combination, as we have done here in the modus ponens that gives us step 3, we do so in the subproof. This is indicated by the fact that the number of the later hypothesis appears in the index of the conclusion.

Indices are important for tracking which conclusions depend on which hypotheses. They help to eliminate non-sequiturs, like the one presented in the introductory section above, and they guide the way in which premises can be discharged in conditional proofs, which are the topic of the next section.

1.4 Implication

In natural deduction systems we do not usually merely display proofs with hypotheses. We discharge premises to prove theorems of a system. The key rule that we will use here is the rule of conditional proof, or $\rightarrow I$ (implication introduction), viz.,

From a proof of B_α from the hypothesis $A_{\{k\}}$ to infer $A \rightarrow B_{\alpha-\{k\}}$,
where k occurs in α .

The proviso that k occur in α is essential. It ensures that, in this case, A is really used in the derivation of B . And so relevant implication also captures this notion of the real use of a hypothesis in bringing about a conclusion. The

³ Readers familiar with E. J. Lemmon's introductory logic book (Lemmon 1965), will recognise this style of proof. Lemmon's system, however, is set up for classical logic rather than relevant logic. One main difference is that the two systems have different conjunction introduction rules.

reason that this rule is key in proving theorems is that it allows us to remove numbers from an index. In the terminology of natural deduction, it allows us to discharge hypotheses.

This rule of implication introduction forces the antecedent and consequent of a provable implication to have something to do with one another. Suppose that $A \rightarrow B$ is provable in this system. Then, it can be shown that A and B have some ‘non-logical content’ in common. That is, they share at least one propositional variable. This is called the ‘variable-sharing property’ or the ‘relevance property’.

In previous sections, we have been discussing the problem of logical relevance in terms of inference, but we can also think about it in terms of implication. Relevant logic was developed in part to avoid the so-called paradoxes of material and strict implication. These are formulae that are theorems of classical logic, but are counterintuitive when we think of the arrow as meaning ‘implication’ in any ordinary sense of the term, or any pre-logical philosophical sense.

Let us look first at the paradoxes of material implication. Among them are the following:

M1 $A \rightarrow (B \rightarrow A)$ (positive paradox);

M2 $\sim A \rightarrow (A \rightarrow B)$;

M3 $(A \rightarrow B) \vee (B \rightarrow A)$;

M4 $(A \rightarrow B) \vee (B \rightarrow C)$.

These schemes are paradoxical because they seem so very counterintuitive. For example, take paradox M3. If we accept the scheme $(A \rightarrow B) \vee (B \rightarrow A)$, then we are committed to holding that for any two propositions one implies the other. But this violates our preformal notion of implication. The problem with material implication *as a way of representing our pre-theoretical notion of implication* is that it is truth functional. The two truth values by themselves are not enough to determine when an implication holds. Something else is needed.

Could this something else be necessity? That is, should we follow C. I. Lewis in claiming that implication is really strict implication? It would seem not, for strict implication has its own menu of paradoxes. Among these are:

S1 $A \rightarrow (B \rightarrow B)$;

S2 $A \rightarrow (B \vee \sim B)$;

S3 $(A \wedge \sim A) \rightarrow B$ (ex falso quodlibet).

S1 and S2 are instances of the general rule regarding strict implication that all valid formulae follow from any proposition. But according to our pre-theoretical logical intuitions, it does not seem that we are justified in inferring any logical truth from any proposition. Similarly, it does not seem that we are justified in

inferring any proposition from any logical falsehood. But this is exactly what *ex falso quodlibet* says.

Lewis and C. H. Langford give an ‘independent argument’ for *ex falso quodlibet*. Here I reproduce their argument with only minor notational changes.

From a proposition of the form $A \wedge \sim A$, any proposition whatever, B , may be deduced as follows:

Assume $A \wedge \sim A$ (1)

(1) $\rightarrow A$ (2)

If A is true and A is false, then A is true.

(1) $\rightarrow \sim A$ (3)

If A is true and A is false, then A is false.

(2) $\rightarrow (A \vee B)$ (4)

If, by (2), A is true, then at least one of the two, A and B , is true.

((3) $\wedge$ (4)) $\rightarrow B$

If, by (3), A is false; and, by (4), at least one of the two, A and B , is true; then B must be true. ((Lewis and Langford 1959), p. 250)

Is this argument good? I think not.

It seems that Lewis had already seen what was wrong with the argument in his 1917 article ‘The Issues Concerning Material Implication’ (Lewis 1917). There he sets out a dialogue between two characters – X and himself (L). Here is a relevant part of that dialogue:

L. But tell me: do you admit that ‘Socrates was a solar myth’ materially implies $2 + 2 = 5$?

X. Yes; but only because Socrates was not a solar myth.

L. Quite so. But if Socrates were a solar myth, would it be true that $2 + 2 = 5$? If you granted some paradoxer his assumption that Socrates was a solar myth, would you feel constrained to go on and grant that $2 + 2 = 5$?

X. I suppose you mean to harp on ‘irrelevant’ some more. ((Lewis 1917), p. 355)

Let’s apply this line of reasoning to the Lewis and Langford argument given above. Suppose that we ‘grant some paradoxer his assumption’ that $A \wedge \sim A$ is true. We then have to take seriously what would happen in a context in which a contradiction is true. We will explore this issue in depth in chapters 5 and 10 below. We will argue in chapter 10 that it is the last step – the use of disjunctive syllogism – that is unwarranted. If we allow that a proposition and its negation can be both true in the same context, then we cannot infer from $\sim A$ and $A \vee B$ that B is true. For it might be that $A \vee B$ is made true by the fact that A is true alone, despite the fact that B fails to be true.

We will leave topics concerning negation to later. To understand fully the treatment of negation in relevant logic, we need to understand its semantics. Other paradoxes, such as positive paradox are treated by the relevant logician’s insistence that the hypotheses in a proof really be used in that proof. Consider

the following attempt at a proof of positive paradox:

- | | | |
|----|---|------------------------|
| 1. | $A_{\{1\}}$ | <i>hyp</i> |
| 2. | $B_{\{2\}}$ | <i>hyp</i> |
| 3. | $A_{\{1\}}$ | 1, <i>reit.</i> |
| 4. | $B \rightarrow A_{\{1\}}$ | 2 – 3, $\rightarrow I$ |
| 5. | $A \rightarrow (B \rightarrow A)_{\emptyset}$ | 1 – 4, $\rightarrow I$ |

The illegitimate move here is the use of implication introduction in the fourth step. 2 does not belong to $\{1\}$ and so we cannot discharge the second hypothesis here.

1.5 Implication and entailment

In the preceding sections, we have been discussing the notion of implication. As we have said, the relation of implication holds between two propositions if the consequent follows from the antecedent. By itself, this statement does not give us a very good interpretation of implication. It is rather ambiguous. As we shall see throughout this book, there are many different ways in which propositions can follow from one another. The matter is made worse by the fact that various philosophers use the word ‘implication’ to refer to different relations.

In chapter 3 below, we will make our use of ‘implication’ quite precise. But for the moment, we will note that the relation of implication, as we use this term here, is a *contingent* relation. What we mean by a contingent relation can be understood most clearly if it is contrasted to the corresponding necessary relation, which is known as *entailment*. Consider the scheme,

The proposition p follows from $p \wedge q$.

The notion of ‘following from’ here can be taken to be a very strong relationship.⁴ For we can derive by logical means alone the proposition p from the conjunction $p \wedge q$. The connection between p and $p \wedge q$ is a necessary connection. Now consider the following statement:

A violation of New Zealand law follows from not paying income tax on honoraria given for presenting seminars at other universities.

Here the relationship between not paying income tax on an honorarium and a violation of the New Zealand tax code is not a necessary connection. We can easily imagine a world in which the tax code were different such that it made honoraria tax exempt income. When we make claims like the one above, we do so assuming other facts that connect the failure to pay tax and a violation

⁴ I use the phrase ‘can be taken to be’ here because if a strong relationship holds, then every weaker relation holds as well.

of the law, that is, particular facts about the New Zealand tax code. Thus, the way in which the violation of the law and the failure to pay tax are connected is contingent. We say that the relationship between the axioms of arithmetic and $2 + 2 = 4$ is that of *entailment* and the relationship between the failure to pay tax and the violation of the law is that of *implication*.

Like implication, entailment is a relevant conditional. Some philosophers have held that entailment is just necessary material implication. That is, the proposition that A entails B is interpreted as $\Box(A \supset B)$, where the box means 'it is necessity that' and the hook is material implication. But on this analysis every proposition entails every truth-table tautology and every proposition is entailed by every contradiction. These are no less counterintuitive in the case of entailment than they were in the case of implication.

In chapter 3 below, we develop an interpretation of implication that makes sense of the notion of propositions following one another contingently. The theory is called the theory of *situated inference*. The theory gives an interpretation of relevant implication in terms of inferential connections between parts of the world. Very briefly, it says that an implication, $A \rightarrow B$, is true in a part of the world if there is information in that situation that tells us that if A is true in some part of the world, then B is also true in some part of the world. These parts of the world are *situations*. I shall show how the various elements of the natural deduction system for relevant logic can be interpreted in this situational framework.

In chapter 6, we will use the theory of situated inference as a basis for a theory of entailment. I follow Anderson, Belnap and many others in holding that entailment incorporates implication and necessity. Thus, when we add modality to our semantics, we get a semantics for entailment as well. As we shall see, the issues involved are more complicated than this, but that is the essence of the view.

1.6 Relevance and conditionals

Although they are important notions, implication and entailment are related to a class of connectives that are much more prevalent in everyday discourse. These are the natural language conditionals. Natural language conditionals have very different properties from implication and entailment. Implication and entailment, for example, are transitive. Consider the following proof:

1. $A \rightarrow B_{\{1\}}$ *hyp*
2. $B \rightarrow C_{\{2\}}$ *hyp*
3. $A_{\{3\}}$ *hyp*
4. $B_{\{1,3\}}$ 1, 3, $\rightarrow E$
5. $C_{\{1,2,3\}}$ 2, 4, $\rightarrow E$
6. $A \rightarrow C_{\{1,2\}}$ 3 – 5, $\rightarrow I$

Thus, from the premise that $A \rightarrow B$ and the premise that $B \rightarrow C$, we can infer that $A \rightarrow C$. A similar proof is valid for relevant entailment as well.

Our ordinary conditionals, however, are not transitive. Suppose that we have a friend, Nick, who is in poor financial shape, but wants to buy a new boat. You say to me, correctly,

If Nick buys a new boat, he will be completely broke.

And I say, also correctly,

If Nick wins the lottery, he will buy a new boat.

If transitivity were to hold of conditionals, we would be able to infer:

If Nick wins the lottery, he will be completely broke,

which is ridiculous. In chapters 7 and 8, however, I show that our semantics for implication can be used as a framework for a theory of ordinary language conditionals.

But ordinary conditionals are very much like implication and entailment, since they too are relevant. To see this point, let us look at a more traditional view of conditionals. In chapter 7 below, we will discuss the most prominent theories of conditionals at length, but we do not need that sort of detail at this stage. Here let us merely present and attack a straw position, one that is probably no longer held by anyone in the field (at least without a good deal of supplementation), but is very close to the position of many. I will call it 'the standard view'.

I follow the standard view in accepting that we can distinguish between counterfactual conditionals and indicative conditionals. The easiest way to make this distinction is through a fine old example due to Ernest Adams. Consider the following statement:

If Oswald did not shoot Kennedy then someone else did.

This statement is true. We know that Kennedy was shot by someone, so if not Oswald, then it must have been someone else. In addition, someone uttering this conditional does not (at least for the sake of the conversation) assume that the antecedent is true or assume that it is false, i.e. that Oswald did shoot Kennedy or that he did not. The above conditional is an indicative conditional. On the other hand, consider the following conditional:

If Oswald hadn't shot Kennedy someone else would have.

This is a counterfactual conditional. Someone uttering it would assume for the sake of the conversation at least that the antecedent is false. Thus the antecedent of a counterfactual conditional is a hypothesis that is (usually at least) assumed to be false. The standard view holds that indicative conditionals are material

conditionals and that counterfactual conditionals are ‘worlds conditionals’ (to use Frank Jackson’s phrase). In its simplest form (due to Robert Stalnaker), this view holds that a counterfactual conditional ‘If it were the case that *A* it would be the case that *B*’ is true if and only if in the closest possible world to our own in which *A* is true, *B* is true as well.

Let us begin by criticising the standard theory of counterfactuals, because that criticism is slightly easier to see, and then we will move on to the indicative conditional. The standard view does not have an adequate treatment of counterpossible conditionals. On the standard view, a counterfactual conditional is true if and only if at the closest possible world in which the antecedent is true the consequent is also true. But consider the following counterfactual:

If Sally were to square the circle, we would be surprised. (1.1)

There is no possible world at which Sally squares the circle. Stalnaker gives all such counterpossibles the value true and we will have the standard view do so too. In respect to (1.1), the standard view seems to get things right, since, intuitively, it is a true conditional. But now consider a closely related conditional, (1.2) below.

If Sally were to square the circle, we would not be surprised. (1.2)

On the standard theory, (1.2) is also true. But this seems to get things wrong, at least according to our intuitions on the matter. Let us call a pair of counterpossibles such that one differs from the other only by virtue of a negation in its consequent and such that they have different truth values, *non-trivial counterpossible pairs*.

These are not isolated examples. There are many cases of non-trivial counterpossible pairs. For instance,

If water were an element, it could not be decomposed into hydrogen and oxygen.

This statement is true, and its counterpossible mate is false. Likewise, the following counterpossible taken from Hartry Field is true, but its mate is false:

If the axiom of choice were false, the well-ordering theorem would fail. ((Field 1989), p. 237)

The antecedent of this conditional is impossible, if the axiom of choice is true.⁵ Then, the counterpossible ‘If the axiom of choice were true, the well-ordering

⁵ Field does not think that the antecedent of this counterfactual is impossible. He thinks that there is no most general notion of possibility. In other words, he rejects the notion of metaphysical possibility. I, on the other hand, will dogmatically accept metaphysical possibility without argument.

postulate would hold' would be true and its counterpossible mate would be false.

Our counterpossibles have all so far had antecedents that are metaphysically impossible. I follow the mainstream in the metaphysics of modality in equating metaphysical and logical possibility. But many philosophers seem not to have adopted this position. For them, let us consider a few straightforwardly 'counter-logical-possibles'. First, it would seem that it is true that,

If Sally were to prove A , we would be surprised,

where A is some long truth-functional contradiction. For an even more blatant example, consider the following statement:

If ' $p \wedge \sim p$ ' were true, then there would be a true contradiction.

Its mate too is false.

Similar examples can be found for indicative conditionals. Suppose that you have made a bet with me that our friend Sally cannot square the circle. Also suppose that she has been given a time limit and that limit has been reached, but neither you nor I know of her success. Thus, it would seem that,

If Sally has squared the circle, you lose the bet

is true and,

If Sally has squared the circle, you win the bet

is false. If we take the indicative conditional to be a material conditional, then both of these statements turn out true. So, just as in the case of counterfactuals, the standard view seems inadequate to deal with indicative conditionals.

The standard view also has difficulties with the failure of an antecedent to be relevant to the consequent of a conditional. To take an example from a poem by the nineteenth-century poet, Frederick Locker-Lampson:⁶

If I pick this guinea pig up by the tail, its eyes will fall out.

No one would believe this conditional merely on the basis of the fact that guinea pigs have no tails. Yet the standard view takes it to be true. Moreover, the counterfactual,

If I were to scare this pregnant guinea pig, its babies would be born without tails.

is counted as true by the standard view, yet it seems false. Relevant logicians, like myself, have placed the blame in both of these cases on a failure on the

⁶ Locker-Lampson's poem is called 'A Garden Lyric'. This example was first used to motivate relevant logic by Alan Anderson (Dunn 1986). I am very grateful to Lou Goble for tracking down the poem. Previously, many had thought that Anderson had invented the example.

part of the standard view to take relevance into account. We claim that there has to be some real relation between the antecedent and consequent of a true conditional. We say that any good analysis of conditionals should incorporate relevance.

In chapters 7 and 8 below, we develop a semantics for indicative and counterfactual conditionals, which uses as a basic framework the semantics for implication developed in chapters 2, 3, and 4. The use of the semantics for implication as a basis for the semantics for conditionals allows conditionals to inherit some of the relevance properties from implication.

1.7 Why are these examples important?

David Dowty, Robert Wall, and Stanley Peters consider the project of semantic theory to be quite similar to the Chomskian project of finding a universal grammar. What a semantics should explain, on their view, are certain sorts of linguistic intuitions held by native speakers. They say

In constructing the semantic component of a grammar, we are attempting to account not for speakers' judgements about grammaticality, grammatical relations, etc., but for their judgements of synonymy, entailment, contradiction, and so on. ((Dowty *et al.* 1985), p. 2)

On this view, the way a semantical theory should be confirmed or refuted is quite straightforward. Semanticists should judge the success of their project against the fit between the entailments, synonymies, contradictions, and so on, that it forces and native speakers' intuitions about these phenomena. This is what I try to do in this book.

Counterpossibles play an important role in explanation and thought. For example, we express the entrenchment of beliefs by means of dispositional counterfactuals. I might express the fact that my belief that my dog Ramsey loves pizza is deeply entrenched by 'If Ramsey were to refuse a piece of pizza, I would be shocked.' Likewise, we can analyse the entrenchment of our beliefs in logical necessities by counterpossibles, such as the Sally sentences given above.

Consider the view of Stalnaker that part of what constitutes adopting a conditional belief is to adopt a doxastic strategy (Stalnaker 1984). If the antecedent comes true, adopt a belief in the consequent. This theory seems right for some cases, if not for all. We adopt the same sort of strategies in some cases concerning counterpossibles. If I believe, say, that $\sqrt{2}$ would be rational if the circle could be squared with a compass and straight edge alone, then I adopt the strategy of believing that $\sqrt{2}$ is rational upon discovering that the circle can be squared. Note that what I do not do is merely adopt a strategy about what

to do if I have a justified belief in the antecedent. I want my beliefs to accord with the truth and so I want to modify them in accordance with what is true, not what merely appears to be true. So the *content* of my belief, it would seem, should reflect what goes on in situations in which the circle really is squared.

One might reply that strategies which depend upon a counterpossible proposition's obtaining are useless. But the fact that the antecedent will in fact never come true is beside the point. The antecedents of most counterfactuals will never come true (after all, counterfactuals are usually stated by people who think their antecedents are false). The point is that the believer may not know this (or even may not be certain of it). We are here not talking about strategies that are supposed to be used but about strategies that give us at least part of the analysis of the contents of beliefs.

With regard to the phenomena of relevance failure, I have similar remarks to make. The desire for a relevant analysis of the conditional seems rather firmly entrenched both in the history of logic and in philosophy of language, as I argue in chapter 7 below. Moreover, our intuitions about relevance with regard to conditionals seem quite robust. Thus, it would seem that only by having an analysis that pays attention to relevance can we do justice to our thinking about conditionals.

As I will argue in chapter 3 below, it is natural to cash out informational dependencies in terms of relevant logic. Taking conditionals to be ways of asserting informational dependencies gives a uniform and natural treatment of these bits of language. I shall argue that this relevant theory accounts for the phenomena of relevance and relevance failure better than its alternatives. It accounts for a wider class of cases and it does it more naturally.

1.8 Why not pragmatics?

One question that I often hear is this: why not start with a possible world semantics and use pragmatics to deal with non-sequiturs, counterpossibles, and the other phenomena that we have discussed? Pragmatics is a branch of linguistics and the philosophy of language that studies features of context. For example, in semantics we treat the word 'but' as being synonymous with the word 'and'. That is, we treat them both as having the same truth table. But they have very different pragmatic features. We use 'but' to indicate that there is some tension between the two statements that we are connecting with it. The word 'and' is much more neutral. This tension is somehow indicated by the use of 'but', although the tension is not logically implied by it. Likewise, one might think that uses of 'implies' have a pragmatic feature that indicates a relevance between antecedent and consequent.

First, I know of no serious, rigorous, attempt to give a pragmatic solution to these problems. Until such a theory is presented, it is difficult to say whether

a pragmatic theory would be successful. Second, theories of pragmatics are notoriously vague. They tell us, for example, to reject the above argument because it violates the Gricean maxim to ‘be relevant’. What counts as relevant is left unsaid in Grice’s theory. Surely, if there is a theory of relevance that is more rigorous than this, it would be better, all things being equal, to appeal to the more rigorous theory. And relevant logic does provide a very specific view about what counts as a relevant deduction. Third, there is a much deeper difficulty with any pragmatic solution. The problem with our proof of Fermat’s Last Theorem is that *the premise does not establish the conclusion*. Surely it is the primary job of logic to give us a theory of what premises establish what conclusions. If our logical system does not do that, it is defective as a theory of ordinary reasoning.

1.9 So what’s the down-side?

So far, the case we have presented for relevant logic seems very strong. Our intuitions support the need for relevance constraints on proof. These constraints, in the form of the requirement that premises really be used in deductions, are easy to formulate and understand, at least for the fragment of the natural deduction system that we have considered so far. If this were all there were to the issue, we could stop here.

Contemporary logic, as done by philosophers, computer scientists, and mathematicians, incorporates two fields – proof theory and model theory.⁷ We have already looked at some of the proof theory of relevant logic. In the next chapter we will turn to model theory and its relationship to proof theory. There we will discuss the reason why model theory has an important place in philosophical logic and what model theory can do for us. Briefly, model theory – also known as ‘formal semantics’ or merely ‘semantics’ – gives a partial specification of a theory of meaning and a theory of truth for a language.

As we shall see, the semantics for relevant logic has both important similarities and differences to the semantics for more widely accepted logics, such as classical logic and modal logic. The differences between the semantics for those logics and that for relevant logic, however, need to be understood and motivated, and this is the task of chapters 3 through 6.

1.10 A note on the metalanguage

Although I am interpreting and defending relevant logic in this book, the logic I use to talk about it is classical. In particular, the logic used to formulate its

⁷ This is not quite accurate. Modern logic is made up of proof theory, model theory, set theory, and recursion theory. But set theory (apart from the brief discussions in chapters 4 and 11 below) and recursion theory are beyond the scope of this book.

semantics is first-order classical logic (i.e. the standard logic taught in introductory logic courses). The use of classical logic in doing mathematics is vindicated in chapter 10. This justification covers the use of classical logic in semantics as well. I use classical logic for two reasons. First, I am trying to defend relevant logic against criticisms levelled largely by philosophers who accept classical logic. Since I accept classical logic as well (in the way explained in chapter 10), classical logic is a language that is common to both myself and the people against whom I am arguing. Second, classical logic is a relatively simple logic and its use makes ideas clearer.

2 Possible worlds and beyond

2.1 Introduction

In this chapter we present the ‘relational’ semantics for relevant logic. We begin by introducing the motivations for truth-theoretic semantics. We also look at the possible world semantics for modal logics. The semantics for relevant logic is a modification of the semantics for modal logics. Looking at the semantics for modal logics will allow us to introduce in an easier form certain ideas that are repeated in the relevant semantics. In addition, as we shall see, the modal semantics has certain important virtues. We shall show, in this and the next chapter, that the relevant semantics has these virtues as well.

2.2 Why truth-theoretic semantics?

In the last chapter we looked at a fragment of the natural deduction system for a relevant logic. We saw that it uses similar rules to those of the natural deduction system for classical logic, but restricts them by forcing the premises of an argument really to be used in that argument. Why not stop there and say that this is all there is to relevant logic? There is something still missing. First, we need to understand the rest of the logic – its treatment of conjunction, disjunction and negation. As we shall see, it is difficult to interpret them entirely in terms of the natural deduction system. Second, what we have said so far is inadequate as an interpretation of implication. We also need some way of understanding, not only how to prove that an implication is true, but what it means for an implication to be true.

Now, we could develop an interpretation of relevant logic in which the theory of truth and meaning of the connectives is explicated by the natural deduction system. This, in effect, is what Michael Dummett has done for intuitionist logic.¹ On Dummett’s interpretation of intuitionist logic, the notion of truth is eliminated and replaced by the notion of constructive proof. An intuitionist does not say that a statement is true, but rather that it can be asserted in the sense that there is adequate evidence to prove it. Now, as Dummett notes, the

¹ Following the work of Jan Brouwer, Arend Heyting, Andrei Komolgorov and Dag Prawitz.

replacement of the notion of truth with the notion of assertibility forces on us the need to recast what we say about various things in a rather radical way. For example, we can no longer say that there may be truths of mathematics that are impossible to prove, for this transcendent notion of truth no longer makes sense.

Or it might seem that we could use the natural deduction system to develop an ‘inferential semantics’ for relevant logic. Wilfred Sellars has distinguished between referential and inferential semantics (Sellars 1963). A referential semantics is the sort of semantics that we will examine throughout most of this book. It takes the meaning of an assertion to be its truth conditions. Moreover, as we shall see soon, the truth conditions of an assertion are determined (at least in part) by the referents of the terms in it. An inferential semantics, on the other hand, takes the meaning of a sentence to be determined by commitments taken on by people who utter that sentence. For example, making an assertion commits one to providing evidence that this assertion is true. In his books *Making it Explicit* (Brandom 1994) and *Articulating Reasons* (Brandom 2000), Robert Brandom develops inferential semantics and defends it. In (Lance and Kremer 1996), Mark Lance and Philip Kremer use Brandom’s view to interpret a natural deduction system for relevant logic as formalising the notion of linguistic commitment. Roughly, Lance and Kremer interpret a sentence $A \rightarrow B$ as saying that anyone who believes A is committed to believing B as well. Although the Lance–Kremer view is interesting and worth pursuing, I will not examine it here. I have two reasons for setting it aside. First, there is a technical point. As Lance has argued, the logic of commitment is not the logic R of relevant implication, but rather the logic E of relevant entailment (Lance 1988). In this book I am particularly interested in interpreting and defending R. Second, there is a strategic point. My strategy in this book is to show philosophers who hold rather conservative views about semantics and metaphysics that they can accept relevant logic without changing their views very much. So I want to avoid radical claims as much as possible.

Thus, we will look at the truth-theoretic semantics for relevant logic. To explain what a truth-theoretic semantics is all about, we will take a quick detour through the history of this field.

2.3 From Tarski

The modern history of truth-theoretic semantics begins in 1932 with the publication of Alfred Tarski’s ‘The Concept of Truth in Formalised Languages’ (Tarski 1985). We won’t go through Tarski’s theory as he presented it, but a standard modernised version of Tarski’s ideas.

Let’s say that we have a language with a finite number of predicates and individual constants (such as proper names). Then, we can construct *recursive*

truth conditions for the sentences of that language. Tarski explicitly developed his theory for formal languages only, not for natural languages, which he thought were logically incoherent. Let's follow Tarski and set out a theory of truth for a simple formal language. Our language contains proper names, predicates, parentheses, and the connectives, $\wedge$ (conjunction), $\vee$ (disjunction), and $\sim$ (negation). Sentences are formed in this language using the standard formation rules. The semantics that we will give to this language is the classical semantics. That is, the logic of this section of the book is a classical logic.

Modernisations of Tarski's 1932 paper make truth relative to a *model*. A model is a mathematical structure. We shall look at a variety of models in this book. Here we will start with a simple model. This model contains a set of individuals, I , called the 'domain', and a function, v , called a 'value assignment'. The value assignment is a function from individual constants to individuals (i.e. to elements of I) and, for all natural numbers n , from n -place predicates to n -ary sequences of elements of I . The model $\mathcal{M}$ is a pair (I, v) , where I is a set of individuals and v is a value assignment. Thus,

- Where P is an n -place predicate and $c_1, \dots, c_n$ are individual constants, $P(c_1 \dots c_n)$ is true on $\mathcal{M}$ if and only if $\langle v(c_1), \dots, v(c_n) \rangle$ is in $v(P)$ (if so, $\langle v(c_1), \dots, v(c_n) \rangle$ is said to 'satisfy P according to v ');²
- Where A and B are formulae of the language, $(A \wedge B)$ is true in $\mathcal{M}$ if and only if both A and B are true in $\mathcal{M}$;
- Where A and B are formulae of the language, $(A \vee B)$ is true in $\mathcal{M}$ if and only if at least one of A and B are true in $\mathcal{M}$;
- Where A is a formula of the language, $\sim A$ is true in $\mathcal{M}$ if and only if A is not true in $\mathcal{M}$.

We can see now how these truth clauses work. The first clause sets out a base case. Suppose that we were to tell a computer the values of v for all of its arguments. Then it would know what sequences of individuals satisfy what predicates according to v . Given this knowledge, the other clauses act as instructions to the computer to generate the truth conditions for any formula of our language.

Around Tarski's theory, a powerful philosophy of language has developed. This philosophy holds that we understand what a sentence means if and only if we understand what would make it true (i.e. its truth conditions). We begin with an understanding of the truth conditions of atomic formulae. For example, I know what it would be for the desk in front of me to be red all over and I know what it would be for my dog to bark, and so on. Luckily for us, it is not the job of logic to explain how we understand what makes atomic formulae true. That is for the philosophies of language and mind (as well as psychology) to explain. As we have seen, given our understanding of the truth conditions for

² where $\langle v(c_1), \dots, v(c_n) \rangle$ is a sequence made up of $v(c_1), \dots, v(c_n)$.

atomic formulae, we can use the recursive truth clauses for the connectives to generate truth conditions for any sentence of our simple language, just as the computer described above could. On the truth-conditional theory of meaning, given our grasp of the truth conditions for atomic formulae and our understanding of the truth-clauses for the various connectives, we can understand any complex sentence of the language. Thus, we get an explanation of how we can come to understand sentences that we have never heard before.

This view of language understanding has shaped a good deal of the history of logic and philosophy of language during the twentieth century. In particular, it has had a very important influence on the development of views about implication. And to these we now turn.

2.4 To possible worlds

As we saw in chapter 1 above, C. I. Lewis proposed his logics of strict implication as ways of capturing the intuitive meaning of ‘implication’. We will examine the semantics of this notion in the present section. This will allow us to see why this semantics cannot capture the notion of relevance that we are after and why our notion needs a more complicated semantics.

Some caveats are in order here. First, we are not going to deal with C. I. Lewis’s preferred logic, S2. That logic has a rather unusual semantics and the point that we wish to make can more easily be demonstrated using normal modal logics. So that we can understand the semantics for normal modal logics, we will set down a few definitions. A modal logic is, for the purposes of this book, a logic that contains the symbols $\Box$ (‘it is necessary that’) and $\Diamond$ (‘it is possible that’). A *classical modal logic* is a logic that contains the scheme $\Diamond A \equiv \sim \Box \sim A$ (which we call ‘Def $\Diamond$ ’) and is closed under that following rule:

Where $A \equiv B$ is a theorem of the logic, so is $\Box A \equiv \Box B$.

A *regular modal logic* is a logic that includes Def $\Diamond$ and is closed under the rule

Where $A \supset B$ is a theorem of the logic, so is $\Box A \supset \Box B$.

And a *normal modal logic* is a logic that is regular and is closed under the rule

Where A is a theorem of the logic, so is $\Box A$.

It is easy to see that every regular and normal modal logic is also classical.

We will look here only at the semantics of normal modal logics. This semantics was developed in the late 1950s (and published in the early 1960s) by Saul Kripke. There is a good deal of debate about who really was the first person to discover the possible world semantics for modal logic. What is clear is that Jaakko Hintikka, Stig Kanger and Arthur Prior, among others, at least

had ideas that were very close to Kripke's. The semantics begins with a set of indices called 'possible worlds'. The notion of a possible world was introduced into philosophy by Leibniz, in the seventeenth and eighteenth centuries. A possible world is a complete universe. In different possible worlds different things happen. For example, in our world there are no talking donkeys, but in other possible worlds there are talking donkeys. There is also a binary relation (a two-place relation) on possible worlds. This is called an 'accessibility relation'.

In possible world semantics, truth is relativised to worlds. So, instead of talking about the truth or falsity of a formula in a model, we talk about the truth or falsity of a formula in a particular world in a model. This relativised notion of truth is used to provide truth conditions for the modal operators, viz.,

$\Box A$ is true in world w in a model $\mathcal{M}$ if and only if for every world w' such that w' is accessible from w , A is true at w'

and

$\Diamond A$ is true in world w in a model $\mathcal{M}$ if and only if there is some world w' such that w' is accessible from w and A is true at w' .

One might wonder what gain is made by translating modal talk into talk about possible worlds. But the virtues of this translation are considerable. First, the semantics is couched in the language of set theory. We talk about sets of points (i.e. possible worlds) and relations between these points. Relations, moreover, can themselves be treated as sets of a sort; in the case of binary relations, as sets of ordered pairs of points. Thus, for example, if the accessibility relation holds between w and w' , then the ordered pair (w, w') is among the members of the set that makes up the accessibility relation.³

The language of set theory is extensional, unlike the language of modality and, more to the point, it is very well understood mathematically. A propositional operator is *intensional* if it does not allow us to replace the formulae within its scope with formulae of the same truth value. For example, both ' $2 + 2 = 4$ ' and 'It is cloudy right now' are true, but 'It is necessary that $2 + 2 = 4$ ' is true and 'It is necessary that it is cloudy right now' is false. Thus, 'it is necessary that' is an intensional operator, and the language of modal logic is said to be an *intensional language*. Intensional operators are difficult to treat mathematically. Set theory, however, has no intensional operators and is mathematically easier to use. Moreover, set theory has been around for a long time and a lot of its mathematical properties are very well understood. The translation of modal logic into a set-theoretic semantics has allowed for the proof of many theorems that would have been difficult to prove otherwise.

³ Ordered pairs are themselves sets. The pair (w, w') can be defined as $\{w, \{w, \{w'\}\}\}$.

Second, the translation of modal logic into possible world semantics has given modal logic a compositional semantics. Like Tarski's theory of truth for propositional logic (that we discussed above), possible world semantics treats the truth conditions of complex sentences as being functionally dependent on the truth conditions of the simple sentences that make them up. Thus, assuming the truth conditional theory of meaning, we can say that a person will understand the statement 'necessarily, Ramsey is a dog' if she understands the truth conditions for 'Ramsey is a dog' and the meaning of 'necessarily'. In this way, possible world semantics fits very well into the philosophy of language that I outlined above.

Third, possible world semantics has the virtue of providing an intuitive and philosophically satisfying interpretation of modality. There are various types of modality. For example, there is nomic necessity. A statement is nomically necessary if it is necessary according to the laws of nature. In order to understand nomic necessity in terms of possible world semantics, we think of the accessibility relation in the following way. A world w' is nomically accessible from w if and only if the laws of nature at w are all obeyed at w' . Let ' $\Box_N$ ' be the nomic necessity operator. Thus the formula $\Box_N A$ is true at w if and only if for every world w' that is nomically accessible from w , A is true at w' . That is a simple example,⁴ and there are some others that are more difficult. It does, however, illustrate well the intuitions that stand behind possible world semantics.

Note that I am not claiming that it is philosophically unproblematic or even intuitive to hold that there are possible worlds other than the one in which we live. Most philosophers have trouble believing in the existence of other worlds, or at least admit that a commitment to them is problematic. I am only claiming that talk about other possible worlds is intelligible and can make modal talk easier to understand.

Fourth, possible world semantics can be used to treat many elements of language other than the operators 'possibly' and 'necessarily'. It can be modified to give a semantics for temporal operators, such as 'it will at some time happen that' and 'it will always happen that'. As we have seen briefly in chapter 1 above and will discuss at greater length in chapter 8 below, it can also be modified to give a semantics for counterfactual conditionals. And there are other uses of possible world semantics. For example, Richard Montague showed that it can be used to give a very realistic semantics for large fragments of natural languages. In science, a theory's ability to treat a wide range of phenomena is a good reason to adopt that theory. Possible world semantics has the virtue of being able to treat a wide range of linguistic phenomena.

⁴ It is not without its philosophical problems as well. For it depends on there being a coherent notion of a law of nature, which some philosophers (such as Nancy Cartwright, Bas van Fraassen and Ronald Giere) doubt.

Thus we can see that possible world semantics has four important theoretical virtues. In this and following chapters, I will argue that the semantics for relevant logic has all these virtues as well.

2.5 Neighbourhood semantics

Before we leave the topic of the semantics for modal logic, we should discuss neighbourhood semantics. This notion is crucial to the argument of chapter 3 below.

Neighbourhood semantics is a sort of possible world semantics but a bit different from the semantics for normal modal logic that we discussed above. Suppose that we want to model a logic that is not regular. Thus, we do not want to make valid the rule that allows us to infer from its being a theorem that $A \supset B$ to its being a theorem that $\Box A \supset \Box B$. A way of doing so was discovered by Dana Scott and developed by Krister Segerberg (Segerberg 1971). We start with a set of possible worlds and a relation N . Unlike the previous semantics, this relation is not a relation between worlds, but a relation between worlds and sets of worlds. If X is a set of worlds and NwX , X is said to be a ‘neighbourhood of w ’.

The truth condition for necessity in this semantics is given in terms of neighbourhoods. A truth set for a formula ‘ A ’ is the set of worlds at which ‘ A ’ is true. We represent the truth set of ‘ A ’ by ‘ $|A|$ ’. Thus, ‘ $\Box A$ ’ is true at w if and only if $Nw|A|$. We can see easily how this semantics allows us to model non-regular logics. Suppose that ‘ $A \supset B$ ’ is a valid formula and that $Nw|A|$. So, $\Box A$ holds at w . There is nothing in this semantics that forces $Nw|B|$ so let’s say that it is false at w . So, ‘ $\Box B$ ’ is not true at w . Thus, ‘ $\Box A \supset \Box B$ ’ is false at w and it is not valid.

In chapter 3 below, we will use neighbourhood semantics to give an interpretation of the relational semantics for relevant logic. But, before we can develop an interpretation of the relevant semantics, we have to understand the basic ideas of that semantics.

2.6 Modelling irrelevance

The question we have before us is how to give a truth conditional semantics for relevant logic. We now have some tools that might seem useful in answering that question, but they aren’t enough.

Suppose that we begin with the idea that implication should be a truth function. A function, recall, is a mathematical device that takes an argument (or sequence of arguments) and returns a value. A truth function is a function on the set of truth values – it takes a truth value (or a sequence of truth values) and returns a truth value. If we have only the truth values T (true) and F (false), to define implication as a truth function we need then to find a function that takes

pairs from the set $\{T, F\}$ and returns either T or F. The obvious way to do this is to get the truth function:

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

This, of course, is the truth table for material implication, and so it makes valid the paradoxes that we are trying to avoid. But, given the constraint that we are trying to construct a truth-functional semantics for implication with only two truth values, this table is the best we can get.

For consider an alternative, for example,

p	q	$p \rightarrow q$
T	T	F
T	F	F
F	T	T
F	F	T

Here we have a table that does not make valid the paradoxical formula, ' $p \rightarrow (q \rightarrow p)$ '. For the case in which ' p ' is true and ' q ' is false. Then ' $q \rightarrow p$ ' is true and so ' $p \rightarrow (q \rightarrow p)$ ' has a truth implying a truth and this gives us the value false. But now our implication is practically useless. For one of the main purposes of having an implication connective is that it licenses modus ponens. That is, if we have a true implication ' $A \rightarrow B$ ' and we know that ' A ' is true, we can infer that ' B ' is true as well. But if we adopt this truth table we are never in this situation. For every implication with a true antecedent is false on this truth table. So, we can never use modus ponens. Other alternative truth tables invalidate other intuitive rules. It would seem that, given the constraint of a two-valued truth table, the best we can do is material implication. So, we should not try to give a two-valued functional semantics for anything that we want to call 'relevant implication'.

We could move to semantics with more than two truth values. It turns out that no finite valued truth table can capture relevant implication. The reasons for this are technical and are treated in (Anderson and Belnap 1975) and (Anderson *et al.* 1992). Instead of taking this route, let's see what happens when we attempt to use a Kripke-style possible world semantics for relevant logic.

Suppose now that we have a class of models that include a set of possible worlds and a single binary accessibility relation. Can we use these to model relevant implication?

The obvious way to model implication in this semantics is to use the following

truth condition: ' $A \rightarrow B$ ' is true at a world w if and only if for every world w' such that w' is accessible from w , either ' A ' is false in w' or ' B ' is true in w' . This truth condition says that an implication is true in a world if and only if the corresponding material implication is true in every accessible world.

This truth condition is a great advance over truth tables. It allows us to avoid paradoxes such as ' $p \rightarrow (q \rightarrow p)$ '. But the modal truth condition is just the truth condition for strict implication, and strict implication comes with its own set of paradoxes, as we have seen in chapter 1 above. For example, we now make valid the formula ' $p \rightarrow (q \rightarrow q)$ '. For choose an arbitrary world w' . In every world either ' q ' is true or it is false, hence this is so in every world accessible from w' , either ' q ' is true or it is false. Thus, ' $q \rightarrow q$ ' is true in w' . Again choose an arbitrary world w . In every world accessible from w , ' $q \rightarrow q$ ' is true (as we have just argued). Thus, in every world either ' p ' is false or ' $q \rightarrow q$ ' is true. Hence, ' $p \rightarrow (q \rightarrow q)$ ' is true in every world.

Other truth conditions using Kripke-style models come aground in different ways. There are such models for logics that are almost relevant – that is, they make valid some but not all of the paradoxes, but none for the sorts of relevant logics that we are studying (see (Anderson *et al.* 1992), § 49.1).

We have seen some ways to model irrelevant logics. Now let us go on to look at ways of modelling relevant logics.

2.7 Modelling relevance

When Richard Routley and Bob Meyer developed the semantics for relevant logic, they did so with the possible world semantics for modal logic in mind. They modified the possible world semantics to fit relevant logic. In doing so, they replaced the binary accessibility relation with a three-place (or 'ternary') accessibility relation. Instead of talking about possible worlds, we will now switch to discussing 'situations'. The indices in the Routley–Meyer semantics are not entirely like possible worlds. We will discuss the distinction between worlds and situations more thoroughly in the next chapter, but we can make a few cursory points about it here.

First, whereas worlds are complete, situations can be incomplete. To use the terminology of Barwise and Perry (Barwise and Perry 1983), worlds decide every issue. That is, they tell us, for any proposition, whether that proposition is true or false. Situations, on the other hand, do not decide every issue. In some situations, the information whether a given proposition is true or false is lacking. This property of situations is sometimes expressed by saying that at some situations 'the principle of bivalence fails'. Second, situations need not be consistent. That is, there are some situations that make contradictions true. Possible worlds, on the other hand, are completely consistent.

Routley and Meyer place a three-place (or ‘ternary’) relation on situations. We will follow them and use the letter R for this relation. The Routley–Meyer truth condition for implication is

‘ $A \rightarrow B$ ’ is true at a situation s if and only if for all situations x and y if $Rsxy$ and ‘ A ’ is true at x , then ‘ B ’ is true at y .

Chapter 3 below is mostly concerned with the interpretation of this relation and this truth condition. In order for the Routley–Meyer semantics to have the same sort of status as a theory of meaning and a theory of truth as the possible world semantics for modal logic, it needs a reasonable interpretation. But for now, we will leave this semantics as an uninterpreted mathematical theory.

The Routley–Meyer semantics for relevant logic, however, clearly has some of the same virtues as possible world semantics for modal logic. First, it translates an intensional language into the extensional language of set theory. As in the case of modal logic, this translation has the virtue of enabling logicians to prove theorems about the logic that would have been difficult to prove otherwise. One wonderful example of such a theorem is Alasdair Urquhart’s proof that several relevant logics are *undecidable*. This means that there is no foolproof mechanical method that will always tell us whether or not a given formula is a theorem of this logic. Many quantificational logics are undecidable, but these relevant logics are among the very few naturally motivated propositional logics to have this property. To prove this theorem, Urquhart shows how to construct models of the sort described here from mathematical structures called ‘projective spaces’. Projective spaces (of infinite dimension) have the property that there is no general mechanical method for telling whether a given equation can be deduced from the axiom set that determines a given projective space. This property is very close to undecidability. Urquhart uses his construction of models from projective spaces to transfer this property of projective spaces onto the logics characterised by the models.

Second, as in the case of possible world semantics, the Routley–Meyer semantics provides relevant logic with compositional theories of meaning and truth. If we know the truth values of some sentences at the various situations in our model and we know which situations are related to which other situations, then we know the truth values of compound sentences containing these sentences. (It is clear, given what we have already said that this is the case for sentences containing implication as the only connective, but as we shall see it is true of the other connectives as well.)

The Routley–Meyer semantics is also known as the ‘relational semantics’ for relevant logic, and we will often use this name for it. There are semantics that do not use relations. There is Urquhart’s ‘semilattice semantics’ (Urquhart 1972) and Kit Fine’s ‘operational semantics’ (Fine 1974) that use operators on pairs of situations and there is J. M. Dunn’s algebraic semantics

((Anderson and Belnap 1975), and (Anderson *et al.* 1992)). In this book, we will only examine the relational semantics.

2.8 What about valid formulae?

One very interesting aspect of the relational semantics is its treatment of valid formulae. In the possible world semantics for normal modal logic, a formula is valid in a model if and only if it is true in every possible world in that model.⁵ This definition of validity is very intuitive. To be valid is to be true everywhere – it is to be true no matter what. But this standard definition of validity won't work in the Routley–Meyer semantics. As we have seen, Routley and Meyer set up their semantics so that, for any formula, there will be models that contain worlds at which those formulae fail to hold. Thus, if we adopt the standard treatment of validity, we no longer have any valid formulae. So, what happens to the notion of valid formulae in relevant logic?

Before I give Routley and Meyer's solution to this question, let me say that the problem of valid formulae is less severe than it might seem at first. My view⁶ is that logic has become overly concerned with theorems and valid formulae. What is more important is the notion of a valid inference, and this is where the emphasis lies in the current project. The emphasis on theorems, I think, has its origin in the fact that logic in the nineteenth and twentieth centuries was largely taken over by mathematicians who were used to thinking in terms of mathematical theories. A theory is a set of statements or formulae. Thus, one would think that a logic should be treated as a set of formulae too.

This is not to say that all logicians have been of this tradition. Proof theorists, beginning with Gerhard Gentzen, have made inference their main object of study. As Gentzen said, when presenting his systems of natural deduction,

The formalisation of logical deduction, especially as it has been developed by Frege, Russell, and Hilbert, is rather far removed from the forms of deduction used in mathematical proofs . . . In contrast, I intended first to set up a formal system which comes as close as possible to actual reasoning. ((Gentzen 1969), p. 68)

It is the study of the notion of proof, rather than the study of valid formulae, that is central to our understanding of reasoning – everyday reasoning as well as mathematical reasoning.

But we cannot avoid the topic of valid formulae completely, at least not in relevant logic. As we saw in the introductory chapter above, the natural deduction system for relevant logic allows us to prove theorems – i.e. those

⁵ This is now the standard account. Kripke, on the other hand, had a single designated world in each model and a formula was said to be valid if it was true in each designated world in each model (Kripke 1965b).

⁶ Which I owe largely to discussions with my former colleague, Peter Schotch.

formulae that we can prove with the null-set as subscript. So the question becomes the following. In our semantics how can we represent those formulae that can be proven in the natural deduction system?

In the semantics, we distinguish between ‘logical’ and ‘non-logical’ situations. A formula is *valid in a model*⁷ if and only if it holds at all the logical situations. We shall examine the formal properties of logical situations in chapter 3 below, and there I will give a philosophical explanation of this distinction.

The role of logical situations in our models is rather interesting. To understand this role, we need to understand the notion of a deduction’s being valid in a model. Let’s say that we have a model $\mathcal{M}$. This model includes a set of situations, among which are some logical situations, a ternary accessibility relation, and a value assignment that gives propositional variables the value true at some situations. As we shall see throughout the rest of this chapter there are further constraints on models, but this will do for now. A one-premise deduction from A to B is said to be valid on $\mathcal{M}$ if and only if for every situation s in $\mathcal{M}$, if A is true at s according to $\mathcal{M}$, then B is also true at s according to $\mathcal{M}$.

As we shall prove later, logical situations have an elegant relationship to valid deductions. This relationship is captured by the semantic entailment theorem:

Theorem 1 (Semantic Entailment) *If, the deduction from A to B is valid in M , then ‘ $A \rightarrow B$ ’ is a valid formula in M .*

If a formula is valid in all models, then it is said to be ‘semantically valid’ or merely ‘valid’. Thus, if a deduction is valid in all models, then the corresponding implication is also valid in all models. This relationship between deductions and implications is a semantic version of what is usually called the ‘deduction theorem’. The deduction theorem states that if a deduction is valid (on a proof theory) then the corresponding implication is a theorem of that proof theory.

2.9 *Persistence

(This and the next section below are more technical than the others and can be skipped without dire consequences, but is accessible to and should be of interest to readers with moderate formal training.)

Let’s make matters a bit more rigorous. In order to prove the semantic entailment theorem, we need first to understand a closely related notion, that of the *persistence* or *hereditariness* relation. This is a binary relation on situation, written ‘ $\sqsubseteq$ ’. In chapter 4 below, we will see how to construct this relation. For

⁷ ‘Validity in a model’ is not universally accepted terminology. Sometimes a formula is said to be ‘verified in a model’ instead.

any situations s and t , if $s \leq t$, we say that t *extends* s . The persistence relation is reflexive, transitive and anti-symmetrical. That is, it is a partial order.

The persistence relation can be defined as follows. For any situations t and u ,

$t \leq u$ if and only if there is some logical situation s such that $Rstu$.

We place two constraints on models in order to get the persistence lemma stated below. First, we constrain the accessibility relation such that if $s' \leq s$ and $Rstu$, then $Rs'tu$. Second, we place a constraint on our value assignment, that is, for any propositional variable p , if p is true according to assignment v at s and $s \leq t$, then p is also true at t according to u . These two constraints make it easy to prove our persistence lemma, namely,

Lemma 2 (Persistence) *For any formula 'A', value assignment v , and situations s and t , if 'A' is true at s according to v and $s \leq t$, then 'A' is true at t according to v .*

The proof of this lemma from these constraints is quite easy, and I leave it to the interested reader.

Given the persistence lemma and the definition of the persistence relation, it is quite easy to prove the semantic entailment theorem. First, we take an arbitrary logical situation s . We also suppose that for any situation t and formulae 'A' and 'B', if 'A' is true at t according to v then 'B' is also true at t according to v . We now show that ' $A \rightarrow B$ ' is true at s according to v . We take arbitrary situations t and u such that $Rstu$ and show that, if 'A' is true at t according to v , then 'B' is true at u according to v . By the definition of the persistence relation, $t \leq u$ (since s is a logical situation and $Rstu$). Suppose that 'A' is true at t according to v . Then, by the persistence lemma, 'A' is also true at u according to v . By our assumption, for any situation, if A is true at that situation according to v then so is 'B'. Thus, 'B' is true at s according to v , and this is what we needed to show. For, generalising, ' $A \rightarrow B$ ' is true at s according to v . We generalise again to say that if the deduction from 'A' to 'B' is valid in a model, then ' $A \rightarrow B$ ' is also valid in the model, and that is just the semantic entailment theorem.

2.10 *Modelling inference

So far we have looked at theorems and at inferences with only one premise. In this section, we will discuss what it means for a deduction with an arbitrary number of premises to be valid. The star on the section heading indicates that the material in this section is a bit harder than that in other sections. Readers who know a little more logic than just what is taught in a standard introductory course, however, should be able to understand it.

We have a ternary accessibility relation on situations. We can, however, construct relations of any ‘arity’ greater than three using this ternary relation. An *arity* of a relation is the number of arguments that the relation can take – thus, for example, a ternary relation has an arity of three. We introduce the notation R^n for $n \geq 0$ and give it meaning as follows:

$$\begin{aligned} R^0 st & \text{ if and only if } s \leq t; \\ R^1 stu & \text{ if and only if } Rst; \\ R^{n+1} s_1 \dots s_{n+2} t & \text{ iff } \exists x (R^n s_1 \dots s_{n+1} x \ \& \ R x s_{n+2} t), \end{aligned}$$

for $n \geq 1$. So, for example,

$$R^2 stuv \text{ if and only if } \exists x (Rstx \ \& \ Rxuv),$$

and,

$$R^3 stuvw \text{ if and only if } \exists x (R^2 stux \ \& \ Rxvw),$$

where $\exists x (R^2 stux \ \& \ Rxvw)$ itself simplifies down to $\exists x \exists y (Rsty \ \& \ Ryux \ \& \ Rxvw)$.

Now we can define the validity in a Routley–Meyer model of an inference. We say that an inference from the premises $A_1, \dots, A_n$ to the conclusion B is *relevantly valid* in a model $\mathcal{M}$ if and only if for all situations $s_1, \dots, s_n$ if for each i such that $1 \leq i \leq n$, A_i is true at s_i in $\mathcal{M}$ then, for any situation t such that $R^n s_1 \dots s_n t$, B is true at t in $\mathcal{M}$. We will deal with the philosophical interpretation of these accessibility relations and relevant inference in chapter 3 below. For now, we will leave the system completely uninterpreted.

In terms of the natural deduction system, we can think of the subscripts as referring to situations. The case where there is only one number in a subscript, in say, $A_{\{1\}}$, is easiest. The structure ‘ $A_{\{1\}}$ ’ tells us that there is some situation (call it ‘ s_1 ’) in which the formula ‘ A ’ is true. Where there is more than one number, the interpretation is slightly trickier. Take, for example, the structure ‘ $A_{\{1,2\}}$ ’. The subscript ‘ $\{1, 2\}$ ’ is taken to refer to some arbitrary situation t such that $R s_1 s_2 t$. Similarly, the subscript ‘ $\{1, 2, 3\}$ ’ is taken to refer to an arbitrary situation t such that $R^2 s_1 s_2 s_3 t$. And so on.

The idea here is to generalise the accessibility relation. As we have seen, ‘ $A \rightarrow B$ ’ is valid in a model if and only if the inference from A to B is truth preserving in every situation in that model. Thus, we have a correlation in our semantics between implication and inference. It makes sense to connect the accessibility relation – the element of the semantics that treats implication – with inference more generally. Thus, we use the technique of relational products to generalise the accessibility relation to relations of arbitrary arity and treat inference in terms of this class of relations. In fact, as Meyer suggests in

(Mares and Meyer 2001), we can omit the superscripts altogether. For, this is the ‘same relation’ construed as holding between different numbers of situations. It has the same function in each of these cases – to provide a semantics for deductive inference. We will follow Meyer in this practice and omit superscripts on the accessibility relation.

2.11 Conjunction and disjunction

The truth condition for conjunction is the usual one, viz.,

‘ $A \wedge B$ ’ is true at s if and only if ‘ A ’ is true at s and ‘ B ’ is true at s .

The truth condition for disjunction is also the standard one adopted from possible world semantics. That is,

‘ $A \vee B$ ’ is true at s if and only if ‘ A ’ is true at s or ‘ B ’ is true at s .

The truth condition for conjunction makes valid the following rules in the natural deduction system:

From A_α and B_α to infer $A \wedge B_\alpha$ ($\wedge I$);

From $A \wedge B_\alpha$ to infer A_α or B_α ($\wedge E$).

These rules merely say that if we know that two formulae are true in the same situation, then we can infer that their conjunction is also true in that situation and that if we know that a conjunction is true in a situation, then we can infer that either conjunct is also true in that situation. This very closely corresponds to the truth condition for conjunction.

Whereas there is not much to say about the truth condition for disjunction, there is a lot to say about the disjunction rules in the natural deduction system. Anderson and Belnap’s rules of disjunction introduction are the following:

From A_α to infer $A \vee B_\alpha$

and

From B_α to infer $A \vee B_\alpha$.

These rules are fairly standard. And we can see that they correspond closely with the truth condition for disjunction. They say if a formula is true at a situation then so is the disjunction of that formula and any other formula. What is more interesting is their rule of disjunction elimination. The Anderson and Belnap rule is

From $A \vee B_\alpha$, $A \rightarrow C_\beta$ and $B \rightarrow C_\beta$ to infer $C_{\alpha \cup \beta}$.

Anderson and Belnap's rule can be understood as a relevant version of the disjunction elimination rule used in Gerhard Gentzen and Dag Prawitz's natural deduction systems for intuitionist and classical logics, viz.,

$$\frac{\begin{array}{c} [A] \quad [B] \\ \vdots \quad \vdots \\ A \vee B \quad C \quad C \end{array}}{C}$$

The notation here is easy to understand. A formula in square brackets indicates a hypothesis. Thus, if we have a disjunction and from the hypothesis of each disjunct we can prove a formula C , then we can infer C . This, in effect, is what we have in Anderson and Belnap's inference rule.

Recently Ross Brady has come up with an alternative rule of disjunction introduction. The problem with the Anderson and Belnap rule is that, together with the rest of the introduction and elimination rules, it does not allow the derivation of the principle of distribution, that is,

$$\text{From } A \wedge (B \vee C)_\alpha \text{ to infer } (A \wedge B) \vee (A \wedge C)_\alpha.$$

Clearly, distribution is valid on our semantics. Thus, we need it in our natural deduction system. Anderson and Belnap add it as one of the basic rules of the system. But Brady's new rule allows us to derive distribution.

To do so, Brady adds a new structure type to his system – lines of proof. His new rule of disjunction elimination is the following:

$$(\text{Brady } \vee E) \text{ From } A \vee B_\alpha \text{ to infer } A_\alpha, B_\alpha.$$

The two subscripted formulae separated by a comma constitute a single step in a proof. They introduce two lines of proof. Within a line of proof, we can use all of the standard rules. Here is how we can use this device to prove distribution:

1. $A \wedge (B \vee C)_{\{1\}}$	<i>hyp</i>
2. $A_{\{1\}}$	1, $\wedge E$
3. $B \vee C_{\{1\}}$	1, $\wedge E$
4. $B_{\{1\}}, C_{\{1\}}$	3, Brady $\vee E$
5. $A \wedge B_{\{1\}}, C_{\{1\}}$	2, 4, $\wedge I$
6. $A \wedge B_{\{1\}}, A \wedge C_{\{1\}}$	2, 5, $\wedge I$
7. $(A \wedge B) \vee (A \wedge C)_{\{1\}}, A \wedge C_{\{1\}}$	6, $\vee I$
8. $(A \wedge B) \vee (A \wedge C)_{\{1\}}, (A \wedge B) \vee (A \wedge C)_{\{1\}}$	7, $\vee I$
9. $(A \wedge B) \vee (A \wedge C)_{\{1\}}$	8, $\vee E$

The rule of $(\vee E)$ states that we can infer from C_α, C_α to C_α . That is, we are allowed to conflate two lines of proof if they prove the same thing. Together

Brady's rules of disjunction elimination and comma elimination give us the effect of Gentzen and Prawitz's disjunction elimination rule for classical and intuitionist logic.

Brady's technique of lines of proof is a very elegant addition to the natural deduction system for relevant logic, but since Brady's system is equivalent to Anderson and Belnap's system, and since we will discuss disjunction very little in the rest of this book, we will leave this topic here and I will leave the reader to choose between these systems.

2.12 *Constraints on the accessibility relation

In order to give a semantics for relevant implication and relevant deduction, in the sense determined by the logic R, we need to place a few constraints on the behaviour of the accessibility relation.

As we have said, we interpret the subscript ' $\{1, 2\}$ ' as used in our natural deduction system to represent an arbitrary situation t such that Rs_1s_2t (where '1' refers to s_1 and '2' refers to s_2), and so on, for numbers higher than 2. It does not matter in our natural deduction system in which order we state a hypothesis or in which order we use it in a proof. The order in which information is presented in a deduction is not crucial. We need $\{1, 2\}$ and $\{2, 1\}$ to coincide. To represent the arbitrariness of the order of premises in our semantics we constrain the accessibility relation such that, if Rs_1s_2t , then it is also the case that Rs_2s_1t . More generally, we postulate that for any situations, $s_1, \dots, s_i, s_j, \dots, s_n$ (where n is some number greater than or equal to 2), if $Rs_1 \dots s_i s_j \dots s_n t$ then $Rs_1 \dots s_j s_i \dots s_n t$. In other words, we can permute any two situations except the last one, which stays fixed.

The next constraint is also easily motivated in terms of the natural deduction system. Consider the following deduction:

- | | |
|---|-----------------------|
| 1. $(A \rightarrow B) \wedge A_{\{1\}}$ | <i>hyp</i> |
| 2. $A \rightarrow B_{\{1\}}$ | 1, $\wedge E$ |
| 3. $A_{\{1\}}$ | 1, $\wedge E$ |
| 4. $B_{\{1\}}$ | 2, 3, $\rightarrow E$ |

This is a very simple inference, but it illustrates something interesting. On our interpretation of the subscripts in deductions, $\{1\}$ refers to a situation s . This deduction tells us that if $(A \rightarrow B) \wedge A$ is true in s then B is also true in s . In order to ensure that this argument is valid, we postulate that, for any situation s , $Rsss$. This implies also that for any number of s 's (greater than 2), $Rs \dots s$.

There is one more constraint on the accessibility relation that we need to postulate. This one tells us that we can use the information in a situation as many times as we want in a deduction. It says that if the relation $R \dots s \dots t$ holds (where R is at least of degree 2), then the relation $R \dots s s \dots t$ also holds. For example, if we have $Rstu$ in a model, then we also have $Rsstu$ and $Rsttu$. (But

we do not necessarily have $Rstu$.) This postulate corresponds to the feature of our natural deduction system that we can use any single hypothesis as many times as we want in a given deduction.

2.13 The logic R^+

Now we have enough information to give a definition of a Routley–Meyer model. A *positive relational frame* is a triple $\langle sit, logical, R \rangle$ such that sit is a non-empty set (of situations), $logical$ (the set of logical situations) is a non-empty subset of sit , and R satisfies the following postulates:

- If $R \dots st \dots u$, then $R \dots ts \dots u$ (interchange);
- If $R \dots s \dots t$, then $R \dots ss \dots t$ (repetition);
- $Rsss$ (complete reflexivity);
- If $Rstu$ and $s' \sqsubseteq s$, then $Rs'tu$,

where $\sqsubseteq$ is defined as follows:

$t \sqsubseteq u$ if and only if there is some logical situation s such that $Rstu$.

A *positive relational model* is quadruple $\langle sit, logical, R, v \rangle$ such that $\langle sit, logical, R \rangle$ is a positive relational frame and v is a value assignment, i.e. a function from propositional variables to sets of situations such that:

- For any value assignment v , any propositional variable p , and any situations s and t , if $s \in v(p)$ (i.e. if p is true at s according to v) and $s \sqsubseteq t$, then $t \in v(p)$.

The truth clauses for formulae can now be stated. We use the notation ' $s \models_v A$ ' to mean that A is true in s according to v . For any propositional variable p and any formulae A and B , we have:

1. $s \models_v p$ if and only if $s \in v(p)$;
2. $s \models_v A \wedge B$ if and only if $s \models_v A$ and $s \models_v B$;
3. $s \models_v A \vee B$ if and only if $s \models_v A$ or $s \models_v B$;
4. $s \models_v A \rightarrow B$ if and only if $\forall x \forall y ((Rsx y \ \& \ x \models_v A) \supset y \models_v B)$.

In clause 4, the hook (or horseshoe) expresses material implication. As we said in chapter 1 above, we use a classical metalanguage to describe our logic. This will be justified in chapter 10 below.

Every instance of each of the following schemes is valid in the class of models that we have just described:

Scheme	Name
$A \rightarrow A$	Identity
$(B \rightarrow C) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$	Suffixing
$(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$	Permutation
$(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$	Contraction

Moreover, all logical situations (indeed, all situations) are closed under modus ponens. Hence the valid formulae of our logic are closed under modus ponens.

The logic defined by these theses and the rule of modus ponens is the logic $R_{\rightarrow}$ (the implication fragment of R). This is a system which contains all of the theorems of Anderson and Belnap's system R , which contain only propositional variables, parentheses and the implication connective.

Let's turn now to the valid schemes that also include conjunction and disjunction. The schemes that we make valid can be classed into different groups. First, we have a scheme that roughly corresponds to the conjunction introduction rule in our natural deduction system, viz.,

$$((A \rightarrow B) \wedge (A \rightarrow C)) \rightarrow (A \rightarrow (B \wedge C)).$$

And then there are two schemes that roughly express the conjunction elimination rule:

$$(A \wedge B) \rightarrow A,$$

and,

$$(A \wedge B) \rightarrow B.$$

For disjunction, the converse is true. We have two schemes that capture the disjunction introduction rule, viz.,

$$A \rightarrow (A \vee B),$$

and,

$$B \rightarrow (A \vee B).$$

And we have one scheme that roughly corresponds to the disjunction elimination rule:

$$((A \rightarrow C) \wedge (B \rightarrow C)) \rightarrow ((A \vee B) \rightarrow C).$$

This last scheme may not look like it has anything to do with the elimination rule, since disjunction appears for the first time in its consequent. But, 'permuting' this scheme we obtain the following:

$$(A \vee B) \rightarrow (((A \rightarrow C) \wedge (B \rightarrow C)) \rightarrow C).$$

This principle looks very much like the rule of disjunction elimination.

We need one more axiom scheme and another rule to finish off our presentation of R^+ . The scheme is the distribution of conjunction over disjunction:

$$(A \wedge (B \vee C)) \rightarrow ((A \wedge B) \vee (A \wedge C)).$$

The rule is a rule of adjunction. It tells us that we can infer that $A \wedge B$ is a theorem from the information that A is a theorem and B is a theorem. Together

all of these schemes define the negation-free fragment of the logic R. This logic is called 'R⁺' (for 'positive R').

2.14 The problem of interpreting the semantics for relevant logic

Now we have a semantics for relevant implication, at least for the fragment of it that includes implication, conjunction and disjunction. But, as it stands, our semantics is just a pile of mathematics. To make it into a structure that gives meaning to the connectives and to relevant inferences, we need a philosophical interpretation of this pile of mathematics.

Giving a philosophical interpretation to the Routley–Meyer semantics is the most important philosophical problem facing relevant logicians at the present time. As Graham Priest says,

... if the ternary relation semantics is to justify the fact that some inferences are valid and some are not, then there must be some acceptable account of the connection between the meaning of the relation and the truth conditions of [implications]. ((Priest 2001), p. 198)

Priest points out that we use the semantics to justify the acceptance of some inference rules and to justify the acceptance of some formulae as theorems and to reject other rules and formulae. For this justification to be robust, we need a way of understanding the ternary relation. And this interpretation has to justify our treatments of implication and inference.

In the next chapter we provide such an interpretation.

3 Situating implication

3.1 The problem of implication

We introduced the problem of implication in the previous chapter. The problem is to give an intuitive semantics for relevant implication. In the previous chapter, we looked at the Routley–Meyer semantics for relevant logic, but we did not give it an interpretation. We cannot claim that relevant logic has an intuitive semantics until we do so. That is the task of the present chapter.

We begin with our basic ontology – the list of things that we presuppose in our semantics. The elements of our ontology are situations and possible worlds, as well as individuals and sets. We have met situations and worlds already in this book, but here and in the next chapter we will look at them in much more depth.

After we introduce our ontology, we give an intuitive semantics for implication using neighbourhoods, which were briefly introduced in the previous chapter. The bulk of the chapter shows how, given a few reasonable assumptions, we can view the Routley–Meyer semantics to be just the intuitive semantics in disguise. In addition, given the resulting interpretation of the Routley–Meyer semantics, we can justify the postulates (given in section 2.13) which yield a semantics for the relevant logic R.

3.2 Situations and worlds

We have already been introduced in chapter 2 to the notion of a possible world. Situations are somewhat less familiar objects to philosophers and semanticists and we will begin with them. They were introduced into logic in the early 1960s by Saul Kripke, who used ‘evidential situations’ in his model theory for intuitionist logic (Kripke 1965a). Situations were introduced into linguistic and philosophical semantics by Jon Barwise and John Perry in the early 1980s in their work on ‘situation semantics’. The classic text on situation semantics is their book, *Situations and Attitudes* (Barwise and Perry 1983).

Barwise, Perry and others use situations to model the fact that we express partial information when we communicate. We do not know what is going on in

the whole world, but rather tell people what is going on in our own surroundings. These surroundings are situations. But a situation, as I use the term, need not be an agent's physical location. We can use situations, for example, to model an agent's state of beliefs. Situations are structures of information, rather than physical locations, and can include information from various sources. As we shall see in later chapters, this feature of situations can be very useful.

Possible worlds theorists attempt to capture this partiality of information by taking the meaning of a sentence to be a set of worlds. Thus, for example, if I say 'Ramsey is in my house' I am expressing the set of worlds in which Ramsey is in my house at the current time. These worlds may vary in terms of whether it is raining in Wellington at that time, and so on. Thus, I am not communicating what the weather is when I make this statement (see (Stalnaker 1986)).

But there are important differences between the possible worlds theory and situation theory. Consider the following statement:

The dog barked at a possum yesterday.

Traditional treatments of definite descriptions have us searching for a unique dog as the referent of 'the dog', but of course there are many dogs. Here the phrase 'the dog' refers to a particular dog from a particular situation introduced by the context in which the statement is uttered. If I were to utter 'the dog', the phrase would usually refer to Ramsey, who is my dog. Robin Cooper (Cooper 1993) argues that the statement indicates a situation in which there is one dog – the speaker's dog. This is called the 'resource situation' for the statement. The definite description can be treated as picking out a unique dog, even though there is more than one dog in the world.

Now, I am not claiming that there is no way in which a possible worlds semanticist can treat the above sentence. In fact, there are many such treatments in the literature. Instead, I am merely claiming, following the situation semanticists, that the ubiquity of expressions that clearly depend on restricted parts of the world gives us reason to believe that natural language is built primarily to talk about restricted parts of the world and not talk just about complete possible worlds. Unlike possible worlds semantics, situation semanticists build semantic partiality into their semantics at its very foundation. This seems to them – and to me – to be a more reasonable way to do semantics.

From the standpoint of relevant logic, there is another way in which the partiality of situations is useful. According to possible world semantics every tautology is true in every possible world. On the possible worlds theory of entailment, *A* entails *B* if and only if in every possible world in which '*A*' is true, '*B*' is true as well. Thus, every statement entails every tautology. And, thus, the possible worlds approach condemns us to fallacies of relevance. The situational approach to semantics, on the other hand, does not have this consequence, since not every tautology is true in every situation. And so using situations allows us to avoid these fallacies.

We do not, however, eschew possible worlds from our semantics altogether. In fact, we will need them for our analyses of negation and modality. As far as the current chapter is concerned, we require at least that there is one possible world – the actual world.

In situation semantics, there are two sorts of situations – concrete situations and abstract situations. A concrete situation is just a part of the world. It might be a physically connected situation, such as my study and its contents while I am typing this sentence. Or it might be physically disconnected. It might include my lounge and its contents while I am on the phone with my friend in Canada and her lounge and contents during the same period. Such a situation might be needed to understand the meanings of the statements made during the phone call.

An abstract situation, on the other hand, is an abstract representation of a part of a world. It might be an accurate representation or an inaccurate one. Our formal semantics will use abstract situations. In chapter 4 below we will construct abstract situations and discuss what it is for one to be an accurate representation of a part of a world. But what we are modelling with abstract situations is what goes on in concrete situations. And one sort of process that goes on in concrete situations is what I call ‘situated inference’.¹

3.3 Situated inference

When we make an inference about what is happening in the world, we do so from the limited perspective of some situation. But we do sometimes make claims about other situations, even some that are distant in time and space from our own physical location.

Consider the following argument. *Star X is moving in an ellipse. Therefore, there must be another body near it with a heavy gravitational pull.* Astronomers use their observations of the movement of star *X* to hypothesise that there is another star (or other heavy body) near *X*. The astronomers are located in a situation that includes their labs, photographs of the star, and so on. And they use physical laws, facts about the cameras and telescopes used to make the photographs, and perhaps other information, to make inferences about what is going on in the distant part of the universe that includes star *X* and this other heavy body.

We can understand this inference in situational terms. The astronomers are in a situation, *s*. They assume on the basis of images in telescopes and photographs that there is a star *X*, which is moving in an ellipse. That is, they hypothesise a situation *t* such that in *t* there is a star *X* that is moving in an ellipse. And on the basis of the information in *s* (such as laws of nature, facts about cameras, telescopes, the photographs, etc.), they conclude that there is some situation,

¹ The term ‘situated inference’ is used by Alice ter Meulen in a different way. My apologies for any confusion that this causes, but it is a very good name for the process that I am describing (see (ter Meulen 1995), § 1.3).

u , that obtains in the same world in which there is some heavy body affecting the motion of X .

As we saw in chapter 2, we can think of the numbers in the indices in our natural deduction system as denoting situations. When we write ' $A_{\{1\}}$ ', we are saying that the proposition A holds at situation 1.

When we make a hypothesis in an inference in our natural deduction system, we hypothesise that a proposition is true in a situation and that this situation obtains in a particular world. This world is held fixed for the entire inference. Thus, suppose we have an inference that begins with the following hypotheses:

1. $A_{\{1\}}$ *hyp*
2. $B_{\{2\}}$ *hyp*
- $\vdots$

We read this as saying that we hypothesise that there is a situation in a world in which ' A ' holds and that there is a situation in the same world in which ' B ' holds.²

Indices that include more than one number indicate where the information came from that allows us to deduce the existence of a certain type of situation in a world. So, when we write ' $A_{\{1,2\}}$ ' we are saying that there is some situation in the same world as 1 and 2 in which A holds and that we derived this from the information in 1 and 2. In addition, we really used information from both 1 and 2 to derive that A is true in this situation.

Here is a simple inference – a case of modus ponens:

1. $A \rightarrow B_{\{1\}}$ *hyp*
2. $A_{\{2\}}$ *hyp*
3. $B_{\{1,2\}}$ 1, 2, $\rightarrow E$.

Here we start from the perspective of situation 1. In that situation, we have the information that A implies B . Then we assume that there is a situation 2 in the same world that supports the proposition, A . Thus we can infer that there is a situation in that world in which B obtains.

To see how implication fits into the theory of situational inference, let's take another look at the rule of implication introduction.

From a proof of B_α from the hypothesis $A_{\{k\}}$ to infer $A \rightarrow B_{\alpha-\{k\}}$,
where k occurs in α .

In situational terms, this says that, if we can infer, on the basis of the hypothesis that there is a situation in the world in which A holds, that there is a situation

² This interpretation of the Routley–Meyer semantics builds on work by Alasdair Urquhart in (Urquhart 1972) and John Slaney in (Slaney 1990). Instead of situations, Urquhart uses 'pieces of information' and Slaney uses 'bodies of information'.

in which B holds, then we can infer that there is a situation in which $A \rightarrow B$ holds.

We use this reading of the rule of implication introduction to formulate our truth condition for implication:

$A \rightarrow B$ is true at s if and only if there is information in s such that on the hypothesis that there is some situation t in the same world as s in which A holds, we can legitimately derive that there is some situation u in the same world in which B holds.

3.4 Informational links

Our theory of implication depends on its sometimes being the case that we can infer (under hypothesis) what sorts of situations exist in our world. The information in a situation will sometimes justify such inferences. The sort of information that justifies these inferences, I call ‘informational links’, but are often called ‘informational constraints’ or merely ‘constraints’.

The notion of an information constraint has its roots in Gilbert Ryle’s view of laws of nature as ‘inference tickets’. Ryle says that,

A law is used as, so to speak, an inference ticket (a season ticket) which licences its possessors to move from asserting factual statements to asserting other factual statements. ((Ryle 1949), p. 121)

One of the uses of laws of nature is to provide licence for inferences of certain sorts. If we know that it is a law that every massive object warps space, then from the fact that m has mass, we can infer that m warps space.³

Laws of nature are not the only facts that license inferences. I call any fact that licenses a situated inference an ‘informational link’. Keith Devlin gives the following list of kinds of links:

Constraints may be natural laws, conventions, analytic rules, linguistic rules, empirical, law-like correspondences, or whatever. ((Devlin 1991), p. 12)

Jon Barwise and Jerry Seligman give a similar list:

³ Readers familiar with the literature on relevant logic might wonder how I can appeal to Ryle’s views in support of my interpretation of the logic R , when Anderson and Belnap created the logic T (for ‘ticket entailment’) explicitly to capture the notion of an inference ticket. The idea behind T is to distinguish the way in which inference tickets act in derivations from other kinds of propositions. Only inference tickets can act as major premises in arguments (in particular, as the first premise in an implication elimination). But, as Anderson and Belnap themselves point out, Ryle’s view does not itself entail the structure of inference captured in T ((Anderson and Belnap 1975), p. 41). I think that the ticket analogy also supports treating implications as minor premises and basic facts as major premises. Suppose that I am in a train station a . I then know that if I were to have a ticket from a to b , that I could travel to b . Likewise, if I am in a situation s in which A holds, I know that on the hypothesis that if there is a situation in that world which carries the information that licenses the inference from A to B , that there is a situation in the same world in which B holds.

Some of them are 'nomic' regularities, of the kind studied in the sciences; others, such as those relating a map to mapped terrain, are conventional; and others are of a purely abstract or logical character. ((Barwise and Seligman 1997), p. 9)

Note that neither Devlin nor Barwise and Seligman claim to have a complete list of kinds of informational links.

My use of informational links together with situations has its origin in David Israel and John Perry's theory of information. For them, a link is a piece of information that tells us that 'for every situation of type T , there is one of type T^1 ' ((Israel and Perry 1990), p. 10). We do not need to concern ourselves with situation types here. On our view, a link is information that tells us that if there is a situation in this world in which a proposition A holds, then there is also a situation in the same world in which B holds (for some propositions A and B).

How informational links allow us to make inferences about the nature of other situations is quite simple. Let us take another astronomical example. Suppose that some astronomers observe that a distant star has gone supernova. From this fact, the salient laws of nature, and so on, the astronomers can predict the time at which the radiation from that star will strike various bodies in the universe and so they know facts about situations other than the ones which consist rather narrowly of the star's going supernova.

For yet another example, consider a stop sign at an intersection. Suppose that s is a situation which includes the local council's having made a law about traffic signs. According to s , we can infer from a sign's being at a particular corner and facing a particular direction and the fact that the council erected the sign that motorists travelling in a certain direction should stop at the white line on the road in front of the sign. In this way, conventions help to ground situated inferences.

On my view, informational links are themselves contained as information in situations, and vary from situation to situation. For example, the information that a particular convention is in place may be contained in one situation, but not in another.⁴

To be an informational link a relation needs to be *perfectly reliable*. As Barwise and Seligman argue, causal relations are often not reliable enough to be considered informational links. They use the example of a flashlight with an unreliable connection between the button and the light ((Barwise and Seligman 1997), p. 17). Sometimes pushing the button turns on the light, but sometimes it does not when other factors come into play, such as a wire's coming loose. The problem is that unreliable connections do not warrant deductive inference. At best, they can be used to justify *defeasible* inference.

⁴ This accords with the view of Routley and Meyer themselves who say that 'relative to the laws in s , $Rstu$ means that u is accessible from t ; i.e., if the antecedent of an s -law is realised in t , then its consequent is realised in u ' ((Routley and Meyer 1972), p. 195).

A defeasible inference is one that may not hold if extra information comes to light. Implication, in the sense that we mean it here, is a non-defeasible relation between propositions.

It may be, however, that many of our inferences about other situations are in fact defeasible. In this case, we may take implication to be an idealisation. As we shall see in chapter 7 below, we present a view of natural language conditionals in which conditionals are interpreted to license defeasible inferences about situations and the connections between them.

Let us sum up the theory of situated inference. From the standpoint of a situation we can infer what other sorts of situations are in the same world, or would be in the same world, using the information in our situation (such as the informational links contained in it), plus the basic deductive techniques of our natural deduction system.

Let's consider a brief example. Suppose that we are in a situation s in which there is an instance of a law of nature such that we can infer from A to B . Then, we can write,

$$1. A \rightarrow B_{\{s\}}.$$

Now we can show that $(B \rightarrow C) \rightarrow (A \rightarrow C)$ also holds in s . We start by hypothesising that there is some situation 1 in the same world such that,

$$2. B \rightarrow C_{\{1\}}.$$

Then, to show that there is some situation in the same world in which $A \rightarrow C$ is true, we postulate a situation in which A holds, that is,

$$3. A_{\{2\}}.$$

So, from steps 1 and 3, we obtain,

$$4. B_{\{s,2\}},$$

which means that, on the basis of the information in s and in our hypothesised situation 2, we have inferred the existence of a situation in the same world in which B is true. We use steps 2 and 4 and modus ponens to get,

$$5. C_{\{s,1,2\}}.$$

Now we can put together the formula that we want using conditional proof. From steps 3–5, we discharge our second assumption to get,

$$6. A \rightarrow C_{\{s,1\}}.$$

From steps 2–6, then we discharge our first assumption to obtain,

$$7. (B \rightarrow C) \rightarrow (A \rightarrow C)_{\{s\}}.$$

So, we have proven that $(B \rightarrow C) \rightarrow (A \rightarrow C)$ is true at s , which is what we set out to show.

3.5 Modelling situated inference I: neighbourhood semantics

What inferences are legitimate in a situation are determined by the facts in that situation. But mathematical logicians, like myself, like to provide models of inference that employ extensional entities, such as sets and functions. In this way, we can use ordinary tools of mathematics, such as set theory, to analyse the formal system.

In this chapter, I present one of the central notions of Richard Routley and Robert Meyer's model theory for relevant logic – their ternary accessibility relation – and motivate it. But I do not begin by presenting the relation. Instead, we will start in this section by looking at a function on pairs of situations.

We use a relation I that relates a pair of situations to propositions. We call I , the *implication relation*. A proposition here is a set of situations. In fact, only certain sets of situations will do as propositions – as we shall see in section 4.1 below – but the conditions on propositions do not concern us here. Where A is a formula, $|A|$ is the proposition that it expresses, that is, the set of situations in which ' A ' is true.

Where X is a proposition, $IstX$ if and only if we can legitimately infer from the hypothesis that s and t exist in the same world to the conclusion that there is some situation in the same world in which X obtains. Now we can set the following truth condition for implication:

$$'A \rightarrow B' \text{ is true in } s \text{ iff } \forall x ('A' \text{ is true in } x \supset Istx|B|).$$

This truth condition says that an implication is true at a situation if and only if hypothesising the existence in the same world of a situation which makes the antecedent true allows us to infer the existence of a situation in that world in which the consequent holds.

The modelling of situated inference in terms of a function of this sort is a form of neighbourhood semantics. Neighbourhood semantics for modal logic was explained in section 2.5 above. Here if $Ist|A|$, then we say that $|A|$ is a *neighbourhood of s and t* . As far as I know, Lou Goble was the first person to develop a neighbourhood semantics for relevant logics, but our version of it is slightly different than Goble's.⁵

⁵ Goble's theory is developed in several unpublished manuscripts. I hadn't thought of the current interpretation explicitly in terms of neighbourhoods until I saw Goble's work.

3.6 Adding conjunction and disjunction

As we saw in chapter 2 above, our truth condition for conjunction is the standard one, viz.,

‘ $A \wedge B$ ’ is true at s if and only if ‘ A ’ is true at s and ‘ B ’ is true at s .

This truth condition might seem rather straightforward. But the interaction between it and our neighbourhood semantics for implication will be of some importance to us. The first point that we will need to make is that our truth condition for conjunction entails the following identity:

$$|A \wedge B| = |A| \cap |B|.$$

In words, the proposition expressed by a conjunction is identical to the set-theoretic intersection of the propositions expressed by the conjuncts. This identity seems natural.

It also seems natural to want our semantics to allow the following inference:

$$\frac{\begin{array}{l} \text{‘}A \rightarrow B\text{’ is true at } s \\ \text{‘}A \rightarrow C\text{’ is true at } s \end{array}}{\therefore \text{‘}A \rightarrow (B \wedge C)\text{’ is true at } s}$$

Suppose that striking a dry match with oxygen present implies that it will give off light and striking it with oxygen present also implies that it will give off heat. Then striking a dry match with oxygen present implies that it will give off both light and heat. In addition, every proposed theory of implication that I know of makes this inference valid. So, it would seem that we should constrain our semantics to make it valid as well.

Making this deduction valid in our neighbourhood semantics is quite easy. All we have to do is set the following condition on I :

$$(\text{Conj}) (IstX \ \& \ IstY) \supset Ist(X \cap Y).$$

In words, if, from the hypotheses that s and t hold at the same world, we can infer that there is some situation in X that holds in that world and some situation Y that holds in this world, then we can infer that there is some situation in both X and Y that holds in that world.

Let us relate what we have said about the semantics for implication more explicitly to situated inference. Suppose that,

$$Ist|A|$$

and,

$$Ist|B|.$$

In terms of situated inference, we can say that

$$A_{\{s,t\}}$$

and,

$$B_{\{s,t\}}.$$

By condition (Conj), we know that,

$$Ist(|A| \cap |B|).$$

Thus, by the truth condition for implication, we obtain,

$$Ist|A \wedge B|.$$

Translating this conclusion into the language of situated inference, we have,

$$A \wedge B_{\{s,t\}}.$$

As we have also seen in chapter 2 above, corresponding to the above semantic inference, the natural deduction system has the rule,

$$\text{From } A_\alpha \text{ and } B_\alpha \text{ to infer } A \wedge B_\alpha. (\wedge I)$$

The conjunction elimination rule is the obvious one, viz.,

$$\text{From } A \wedge B_\alpha \text{ to infer } A_\alpha \text{ or } B_\alpha. (\wedge E)$$

The rule of conjunction elimination is straightforward, but there is an issue that we need to discuss regarding the rule of conjunction introduction. The introduction rule says that inferring a conjunction is justified only if we have derived the two conjuncts *with the same index*. This is very important. An alternative conjunction introduction rule that might seem reasonable is the following:

$$\text{From } A_\alpha \text{ and } B_\beta \text{ to infer } A \wedge B_{\alpha \cup \beta}. (\wedge I(a))$$

This rule is unacceptable to relevant logicians. For it allows us to infer paradoxes of material implication. Consider the following argument:

- | | |
|--|------------------------|
| 1. $A_{\{1\}}$ | hyp |
| 2. $B_{\{2\}}$ | hyp |
| 3. $A \wedge B_{\{1,2\}}$ | 1, 2, $\wedge I(a)$ |
| 4. $A_{\{1,2\}}$ | 3, $\wedge E$ |
| 5. $B \rightarrow A_{\{1\}}$ | 2 – 4, $\rightarrow I$ |
| 6. $A \rightarrow (B \rightarrow A)_\emptyset$ | 1 – 5, $\rightarrow I$ |

In effect what happens in this argument, and is allowed more generally if we accept $(\wedge I(a))$, is that we allow the addition of irrelevant premises. For suppose that we have a valid inference from the premises $A^1, \dots, A^n$ to the conclusion B . Then, we can take an arbitrary new premise, A^0 and construct an inference

from $A^0, A^1, \dots, A^n$ to B by conjoining A^0 to any of the other premises using $(\wedge I(a))$ and then eliminating the conjunction as is done in the inference above. Then the derivation of B will seem to depend not just on $A^1, \dots, A^n$ but on A^0 as well.

Happily, the theory of situated inference does not warrant $(\wedge I(a))$. For, what it says is that given a situation in which A holds and a situation in which B holds, we can infer that there is a situation in which both A and B hold. Whereas it may be true of any world that for any two situations s and t in it there is a situation u in it in which all the propositions of each of s and t also hold in u , but the inference to the existence of this sort of situation is hardly justified by the information in every situation. It is too much to ask of all our situations that they contain this much information about the nature of worlds.

As we have seen in chapter 2 above, the truth clause for disjunction is the following:

‘ $A \vee B$ ’ is true in s iff ‘ A ’ is true in s or ‘ B ’ is true in s .

Correspondingly, the proposition $|A \vee B|$ is the union of the propositions $|A|$ and $|B|$.

Interestingly, disjunction does not create the same difficulties in the semantics that conjunction does. Recall that the rule of conjunction of consequents posed a problem for our semantical analysis and required the addition of a postulate governing neighbourhoods. Disjunction is the ‘dual’ connective to conjunction and the dual to the principle of conjunction of consequents is the principles of the elimination of disjunction of antecedents, viz.,

$$\frac{\text{‘}(A \vee B) \rightarrow C\text{’ is true at } s}{\therefore \text{‘}A \rightarrow C\text{’ is true at } s}$$

and,

$$\frac{\text{‘}(A \vee B) \rightarrow C\text{’ is true at } s}{\therefore \text{‘}B \rightarrow C\text{’ is true at } s}.$$

The converse rule is,

$$\frac{\begin{array}{l} \text{‘}A \rightarrow C\text{’ is true at } s \\ \text{‘}B \rightarrow C\text{’ is true at } s \end{array}}{\therefore \text{‘}(A \vee B) \rightarrow C\text{’ is true at } s}.$$

These rules follow directly from the structure of our neighbourhood semantics, without additional postulates. I will leave the proof of this to the very interested reader.

More interesting for us are the rules,

$$\frac{\text{‘}A \rightarrow B\text{’ is true at } s}{\text{‘}A \rightarrow (B \vee C)\text{’ is true at } s}$$

and,

$$\frac{'A \rightarrow C' \text{ is true at } s}{'A \rightarrow (B \vee C)' \text{ is true at } s}.$$

Our semantics also justifies this rule. What we use to justify it is the following principle. *If $IstX$ and X is a subset of Y , then $IstY$.* This principle, in turn, is justified by our interpretation of the relation I . What ' $IstX$ ' says is that s contains the information that if t obtains, then so does some situation in X . But if X is a subset of Y , then every situation in X is also in Y . Thus, if $IstX$ is true, then s contains the information that if t obtains so does some situation in Y . Since X is a subset of $X \cup Y$, then if $IstX$, it should also be the case that $Ist(X \cup Y)$ and $Ist(Y \cup X)$. (This is actually redundant, since $X \cup Y = Y \cup X$.)

3.7 Modelling situated inference II: relational semantics

We can generalise what we have said about conjunction. Suppose, for example, we are in a situation s and we hypothesise that there is a situation t and in it an object i which is uniformly green and that i 's surface occupies a region r in space. Let's call the proposition that i is uniformly green, g , and the proposition that i 's surface occupies r , ' o '. From the standpoint of s , $g \wedge o$ implies every proposition of the form r' is occupied by a uniformly green surface, where r' is a subregion of r . According to what we said in the preceding section, the conjunction of pairs of these propositions are also implied by $g \wedge o$ (and conjunctions of these conjunctions, and so on).

It seems reasonable, however, to say that not only does $g \wedge o$ imply each of these propositions and finite conjunctions of these propositions, but it also implies the conglomerate of all these propositions. We only have a finite conjunction in our language, but we should think of this in conceptual, rather than linguistic terms. When we think of what $g \wedge o$ implies, it is intuitive to say that it implies that there is a situation in the present world in which every subregion of r is uniformly green.

We can capture this intuition in our formal semantics by postulating that, for all situations s and t , the intersection of all the neighbourhoods of s and t is also a neighbourhood of s and t .⁶ Let us call the intersection of the neighbourhoods of s and t , $P(s, t)$.

For a moment, let's return to what we said about finite conjunction. As we said, we need to close the neighbourhoods of a pair of situations under

⁶ In order to set the stage for the argument that follows, we have to assume also that the intersection of the neighbourhoods of s and t contains at least one situation.

intersection. From the standpoint of our interpretation of relevant implication, this says something rather interesting. Suppose that X and Y are propositions and X is a neighbourhood of s and t and so is Y . Then $X \cap Y$ would also be a neighbourhood of s and t . This tells us that if s and t obtain in the same world, then so does some situation from $X \cap Y$. We need no extra information to know that some situation obtains in that world in which X holds and some situation obtains in which Y holds. Thus, $IstX$ provides superfluous information as does $IstY$.

The same holds for generalised intersection. Knowing that some situation from the intersection of all the neighbourhood of s and t is in that neighbourhood gives us more information than does the information that any other proposition is a neighbourhood of s and t . Therefore, we really need to state only that this generalised intersection $P(s,t)$ is such that $IstP(s,t)$.

We now define a ternary relation on situations, R .

$Rstu$ if and only if u is in $P(s,t)$, that is, if and only if u is in the intersection of the neighbourhoods of s and t .

We can read $Rstu$ as saying that, given the hypotheses that s and t obtain in the same world, u is among the candidates that we can infer also to obtain in that world. The situation u will contain all the information that can be inferred from the information in s and t . Given the information in s and t , we cannot deductively infer which of the situations u such that $Rstu$ obtain. But we can know that on the hypothesis that s and t obtain at the same world, at least one such u also obtains at that world.

This relation, R , is the Routley–Meyer accessibility relation. Given this definition, we can derive the following theorem from our neighbourhood semantics.

Theorem 3 *' $A \rightarrow B$ ' is true at s iff $\forall x \forall y ((Rsxy \ \& \ 'A' \text{ is true at } x) \supset 'B' \text{ is true at } y)$*

A proof of this theorem can be found in section 3.12 below. The theorem proves the equivalence that Routley and Meyer use as a truth condition for implication. We can read the truth condition as saying that ' $A \rightarrow B$ ' is true in a situation s if and only if, when we postulate the existence of any situation in the same world in which ' A ' holds, we can infer that there is a situation in that world in which ' B ' holds. And this is just our interpretation of relevant implication of section 3.5 above.

Thus, we have accomplished one of the tasks that we set out in the introductory section for this chapter. We have provided an interpretation of the ternary relation in the semantics for relevant logic. Now we need to go on to interpret the other features of the Routley–Meyer semantics.

3.8 A part-whole relation on situations

One situation might be part of another situation. For example, the situation s that just contains all the information that is present in my study while I am writing this sentence is contained in the situation t that contains all the information present in Wellington while I write this sentence. The persistence relation that we discussed in chapter 2 above is a part-whole relation on situations. Thus, if $s \sqsubseteq t$ we can say that s is part of t .

3.9 The place of theorems in a universe of situations

As we saw in section 2.8 above, a formula is valid in a model in the Routley–Meyer semantics if and only if that formula is true in every logical situation of that model. In section 2.8 we saw that the logical situations have a rather neat relationship to the valid inferences that have only one premise. That is, according to the semantic entailment theorem, if it is valid in a model to infer from ‘ A ’ to ‘ B ’, then ‘ $A \rightarrow B$ ’ is valid in that model. As we saw in section 2.10, this relationship extends very nicely to inferences with more than one premise. I am not going to claim any deep philosophical motivation for the distinction between logical and non-logical situations. Rather I appeal to the elegance of this relationship between logical and non-logical situations as a reason to accept this aspect of the model theory. In scientific investigations elegance counts as a virtue of a theory. I think that the same should be said of semantical theories. All things being equal, we should accept the more elegant semantical theory. Thus, the elegance of the relationship between valid inference, valid formulae, and logical situations gives us some justification to accept the current semantical theory.

3.10 Non-normal situations

Very few of the situations in our model are logical situations. The other situations are, in the parlance of modal logic, *non-normal* situations. That is, a non-normal situation is a situation in which some or all of the laws of logic fail. The concept of a non-normal situation is due to Kripke who uses non-normal worlds in his semantics for non-normal modal logics (Kripke 1965b).

There are, however, interesting differences between Kripke’s non-normal worlds and Routley and Meyer’s non-normal situations. In Kripke’s semantics, the truth conditions for the modal operators at non-normal worlds are not the same as their truth conditions at normal worlds. If, as we have assumed, these truth conditions at least in part constitute the meaning of these operators, then on Kripke’s semantics the meanings of the modal operators are not the same at

normal worlds as they are at non-normal worlds.⁷ This is not the case with the Routley–Meyer semantics. The truth condition for implication, and all the other connectives, is the same throughout the frame. Thus, at least in this regard, the Routley–Meyer semantics does not force connectives to change meaning from situation to situation.

What is also interesting is that in our semantics, it is not the presence of non-normal situations that requires explanation. Rather, it is the selection of certain situations as those that determine the logical truths that needed justification. This inverts the strategy of (Priest 1992). Priest suggests that we treat non-normal situations as the sort of situations that would be described by ‘logic fiction’:

[J]ust as there are possible worlds where the laws of physics are different, so there are possible worlds where the laws of logic are different. Anyone who understands intuitionist logic or quantum logic, for example, has some idea of what things would be like if these were correct (assuming, for the sake or argument, that they are not). Few novelists have (yet) explored the genre of stories about worlds where the laws of logic are different (logic fiction?). But it shouldn’t be too difficult to write interesting such stories. . . . Such stories may bend the mind, but no more so than stories set in worlds with strongly non-Euclidean geometries. (Ibid., p. 292)

Logic fictions are stories in which the laws of logic are different. The possibility of our understanding such stories argues for the existence of situations in which the laws of logic are not those of the actual world.

I will not argue against the notion of logic fictions. We certainly could tell stories using alternative logics. And I won’t argue that we could not construct a set of ‘situations’ to model these stories; in fact, I will follow Priest’s lead in this regard in section 8.10 below.

On the other hand, I will argue against Priest’s use of the notion of logic fiction to interpret his semantics. Priest illustrates his notion of non-normal worlds using a simple and attractive model theory. On this semantics, conjunction and disjunction have their usual truth conditions throughout the frame, but implication is given a different truth condition at normal situations than it is given at non-normal situations (ibid., pp. 293f.). I have already expressed qualms about varying the truth condition of connectives from situation to situation, but my difficulty here is with Priest’s interpretation of his semantics. On his semantics, as we have said, the truth conditions for conjunction and disjunction are the standard ones and they hold throughout his frame. Thus, for example, for all situations s ,

‘ $A \wedge B$ ’ is true at s iff ‘ A ’ is true at s and ‘ B ’ is true at s .

⁷ This is not to say that there is no reasonable reading of Kripke’s model theory. See (Cresswell 1967).

Implications, in contrast are made true by non-normal situations in haphazard ways. Thus, contraposition, modus ponens, and other standard rules fail at non-normal situations. If we can understand fictions in which these rules fail, it does not seem farfetched to claim that we can also understand fictions in which the above rule for conjunction fails.

If we want to include all situations that correspond to possible logic fictions, it would seem that the view we should hold is that all sets of propositions correspond to situations. This, in effect, is the view that Takashi Yagisawa supports (Yagisawa 1988). My criticism of the view in Yagisawa's paper⁸ is that it does not predict the behaviour of connectives in ordinary discourse. As is, it gives rise only to a very trivial theory of semantic entailment. Propositions entail themselves and nothing else. Thus I reject both Priest's view of non-normal situations and Yagisawa's very generous view of worlds.

3.11 An alternative interpretation of the relational semantics

Surprisingly, there are very few interpretations of the Routley–Meyer semantics on the market. One rather elegant interpretation is due to Jon Barwise and has been developed by Greg Restall.

The interpretation employs Jon Barwise's channel theory (Barwise 1993). On Barwise's view, information flows between 'sites' by means of 'channels'. A channel exists between sites when there is a flow of information between those sites. In my lounge, I see a news correspondent from in front of the destroyed World Trade Center in New York. I can see this because there is a channel connecting the distant news correspondent to me. The connection from him to me is made up of wires connected in the right ways, transmitters, satellites, and so on. This physical link gives rise to a channel through which information flows.

Barwise (Barwise 1993) and Restall (Restall 1996) have suggested treating the Routley–Meyer semantics in channel theoretic terms. We should think of situations as being sites of information. One site, for example, contains the news reporter at Ground Zero. Another site contains me and my lounge and television set.

Some situations are also channels of information that connect sites. On their interpretation, we read $Rstu$ as saying that s is a channel between t and u .

This interpretation motivates a slightly weaker logic than R . Here is the problem: on the channel theoretic reading, $Rsss$ says that s is a channel from

⁸ Yagisawa, I think, does not hold this view. He seems to hold that this view is the one that, given David Lewis' motivations for modal realism in (Lewis 1986), Lewis should accept. In other words, he takes it to be a reductio of Lewis' position.

itself to itself. It does not seem intuitive to require of all situations that they be channels from themselves to themselves, at least not if the notion of a channel is a generalisation of cases like that of the link between television cameras and a television.

The existence of this other interpretation, however, poses no threat to my own interpretation. The channel-theoretic reading introduces a different use for a relevant logic than the one that I have in mind (i.e. situated inference).

3.12 A proof of theorem 3

For readers who have a more formal bent, I include a proof of theorem 3. The proof is quite simple. It does, however, require the following two assumptions:

1. The set of neighbourhoods of s and t is closed *upwards under supersets*. This means that if $X \subseteq Y$ and $IstX$, then $IstY$.
2. The set of neighbourhoods of s and t is closed under intersections. In particular, the intersection of the whole neighbourhood of s and t ($P(s, t)$) is a neighbourhood of s and t .

Here is a proof of theorem 3. We need to show that

$$s \models_v A \rightarrow B \text{ if and only if } \forall x \forall y ((Rsxy \ \& \ x \models_v A) \supset y \models_v B),$$

where ' $s \models_v A$ ' means " A is true at s according to the value assignment v '. We will prove this biconditional one direction at a time. First we will assume $s \models_v A \rightarrow B$ and show $\forall x \forall y ((Rsxy \ \& \ x \models_v A) \supset y \models_v B)$. This is rather easy. By the truth condition for implication, we have

$$s \models_v A \rightarrow B \text{ if and only if } \forall x (x \models_v A \supset Isx|B|_v),$$

where $|B|_v$ is the set of situations that make B true according to v . Therefore, from our assumption that $s \models_v A \rightarrow B$ we obtain

$$\forall x (x \models_v A \supset P(s, x) \subseteq |B|_v).$$

Then, by the definition of R , we have

$$\forall x \forall y ((Rsxy \ \& \ x \models_v A) \supset y \models_v B),$$

which is what we set out to prove.

The other direction is only slightly harder. We begin by assuming that

$$\forall x \forall y ((Rsxy \ \& \ x \models_v A) \supset y \models_v B)$$

and proving that $s \models_v A \rightarrow B$. By this assumption, for any situation t in $|A|_v$,

$$P(s, t) \subseteq |B|_v.$$

By assumption 2 above, $Ist(P(s, t))$. Thus, by assumption 1, we have

$$Ist|B|_v.$$

Generalising, by the truth condition for implication, we have

$$s \models_v A \rightarrow B,$$

as required, ending the proof of the theorem.

4 Ontological interlude

4.1 Semantics and metaphysics

In this book I am urging people to accept relevant logic and its semantics. In this chapter we discuss what metaphysical consequences are entailed by accepting the semantics. But before we can get into this issue properly, we should discuss what it means to accept a theory.

Theories can be treated either realistically or anti-realistically. A realist about a subject like semantics, will ask of the theory whether or not it is literally true. If it is found not to be true, then the realist will reject the theory. For a realist, to accept a theory is to believe that it is literally true. Anti-realists, on the other hand, do not use the literal truth of a theory as the criterion of acceptance.

There are various brands of anti-realism. Those that are most relevant to the subject of semantic theories are instrumentalists and fictionalists. Instrumentalists about semantic theories hold that talk about things like possible worlds is useful for understanding our semantic intuitions but are not to be taken literally (see e.g., (Kripke 1972)).

Fictionalists also hold that we should not treat the claims made in a semantic theory as being literally true. Suppose, for example, that we adopt a fictionalist attitude towards the semantics for relevant logic. Then, when we say, for example, 'There are situations that make contradictions true' (as we shall say in chapter 5 below), what we mean is that according to this semantics there are situations that make contradictions true. In addition, the fictionalist about this semantics will also hold that the semantics is an appropriate fiction to use to analyse the meaning and truth value of our statements about implication. The fictionalist will say that a statement ' $A \rightarrow B$ ' is true if and only if according to the semantics for relevant logic it is true (Rosen 1990).

Fictionalists trade in the commitments of a given theory for the difficulty of explaining why the theory is appropriate to give truth conditions for a given sort of discourse. A very interesting example of a defence of a sort of fictionalism is given by Richard Joyce (Joyce 2001). Joyce claims that our moral views are not made true by any class of moral facts, but rather that morality is a convenient fiction. He claims that we evolved to believe in certain moral 'truths' such as that

stealing and unjustified killing are wrong. These beliefs help us coordinate our actions and hence to live in societies. Thus Joyce defends our moral discourse without committing himself to the existence of moral facts.

If we adopt an instrumentalist or fictionalist attitude towards the semantics for relevant logic, then our task is made easier. I think that the arguments of this book go some way to arguing that the semantics for relevant logic or, more accurately the intended model, is an appropriate fiction to use to determine the truth or falsity of implicational statements.

On the other hand, if we adopt a realist view towards the semantics, we need to investigate what commitments it has in terms of what entities it postulates and what else it needs in order to make it true. The remainder of this chapter is presented from a realist's point of view. That is not to say that I am rejecting anti-realism. I give no arguments and take no stand either way. But to take realism seriously, we need to investigate the commitments of the theory.

4.2 The reductive project

One way of evaluating a theory is to do so in terms of W. V. O. Quine's *criterion of ontological commitment*. According to the criterion of ontological commitment, if we accept a theory we have to accept the existence of the entities that the theory requires. We determine what entities the theory requires, on Quine's view, by translating our theory into quantified classical logic and then seeing what statements of the form ' $\exists x Fx$ ' ('there are some F s') are derivable in the theory. If ' $\exists x Fx$ ' is derivable in the theory, then the theory is committed to the existence of F s. We won't give a complete formalisation of our semantical theory here. That won't be necessary. We will be able to determine the ontological commitments of our theory in a semi-formal setting.

So far, we have treated situations and possible worlds as primitives in our ontology. We could continue to do so, but many philosophers would find this dissatisfying. One way in which theories are judged, according to a broadly Quinean philosophy of science, is in terms of what ontological commitments they make. The fewer commitments made by a theory, *all things being equal*, the better the theory is.

The main competitors for the relevant theories of implication and conditionals are semantical theories based on possible worlds. These theories, like our own, accept possible worlds, but they eschew situations. And so it looks like they win in terms of ontological commitment. It would be a great bonus for our own theory if we could show that it has no more in the way of ontological commitments than do these possible world semantics. And this, with certain qualifications, is what I set out to do in this chapter.

My strategy is to show that we only need the following entities for our semantical theory – possible worlds, individuals and sets. Situations are constructed

from these primitives. All of these entities are employed by possible world semantics as well. Therefore, in terms of ontological economy, possible world semantics is no better off than is the semantics for relevant logic.

Well, almost. In constructing my models for relevant logic, I use a non-standard set theory. This is Peter Aczel's non-well-founded set theory. We will examine non-well-founded set theory in some detail in section 4.6 below, but I will give a brief description of it here.

As opposed to standard set theories, non-well-founded set theory allows sets to be members of themselves or to be members of members of themselves, and so on. Aczel proved that, if standard Zermelo–Fraenkel (ZF) set theory is consistent, then so is his non-well-founded set theory. Thus, using his set theory is as 'safe' as using ZF. As Jon Barwise and Laurence Moss have shown, non-well-founded set theory provides for an elegant version of possible world semantics for model logic, and I adapt the Barwise and Moss method here to deal with relevant logic. As we shall argue in section 4.6 below, there are other good reasons to accept non-well-founded set theory.

The use of non-well-founded set theory is not essential to the present project. In section 4.12 below, we shall see that models for relevant logic, which incorporate the same basic ideas, can be constructed using standard set theory. But the model theory that employs standard set theory is not as elegant mathematically, and so I accept the version that uses non-well-founded set theory.

4.3 Possible worlds and individuals

The first two elements in our ontology are possible worlds and individuals. There is some controversy in philosophy of logic over what sorts of things each of these are, so we should say a few words about them here.

We have already discussed the debate between realists and anti-realists with regard to semantics. Among realists about possible worlds, there are two camps – vertebrate realists and ersatzers. Vertebrate realists hold that possible worlds are universes like our own. They believe that there are many parallel universes and each contain individuals standing in relations to one another. The most famous (and perhaps only) vertebrate realist is David Lewis. There are many types of ersatzers, but they all believe that possible worlds are abstract objects or constructions from abstract objects. Some believe that possible worlds are sets of sentences, some believe that they are properties or complex properties, and others believe that they are sets of propositions.

Luckily, we do not have to take a stand in the debate between vertebrate realists and ersatzers. In my theory, possible worlds are primitives – they get no analysis. Now, this does not mean that further analysis of them is impossible. Rather, if there is a way of analysing possible worlds in terms of more primitive (and more satisfactory) entities, then my theory can accept this analysis. We need possible worlds, wherever they come from.

We also need individuals. The two most common theories of individuals on the market at the present time are the theory of counterparts and the theory of trans-world individuals. On counterpart theory, no individual exists in more than one world. On the theory of trans-world individuals, on the other hand, the same entity can exist in many worlds. On this issue, I will side with the theory of trans-world individuals. There are technical reasons for doing so. It is very difficult to give an adequate semantics for a logic using counterpart theory. I won't give a knock-down argument for this claim here, but rather merely point to the history of the use of counterpart theory in modal logic. David Lewis' (Lewis 1968) counterpart-theoretic semantics for modal logic is badly flawed in the sense that it does not characterise any normal modal logic (see (Hazen 1979) and (Woollaston 1994)).¹

4.4 Properties and relations

I take properties and relations to be functions from worlds to sets. What sort of sets are in the range of a property depends upon the type of that property. For example, a monadic property of individuals is a function from worlds to sets of singletons of individuals, a binary relation on individuals is a function from worlds to sets of pairs of individuals, and so on. I follow standard practice in maths and philosophy in taking functions to themselves be sets of sorts. Thus properties and relations, in this sense, are reducible to the primitive elements of my construction.

The types of properties and relations that we will need in future chapters extends beyond first-order properties and relations of possible individuals. We will require a variety of properties of properties and relations between properties, and so on. Like first-order properties and relations, higher-order properties and relations are just functions; in this case they are functions from worlds to sets of n -tuples of properties or relations. Clearly, since we have already assumed a standard set theory, accepting higher order properties of this sort is no extra ontological commitment at all. This commitment is derivable from our basic ontology.

We adopt standard notation from the logical theory of types to be used whenever we need to indicate the type of an entity. Our basic type is that of an individual, ι . We can now define the hierarchy of types recursively. Our set of types is the smallest set such that ι is in it and if $\tau_1, \dots, \tau_n$ are types, then so is $(\tau_1, \dots, \tau_n)$. So, for example, the entity $P^{(\iota)}$ is a property of individuals, $R^{(\iota, \iota)}$ is a binary relation between individuals, $S^{((\iota, \iota), \iota)}$ is a relation between properties of

¹ Graham Forbes avoids the difficulties in Lewis' semantics, but his semantics has two different truth conditions for statements of the form ' $\Box A$ ', depending on the form of ' A ' ((Forbes 1985), p. 63). This sort of technical complication is best avoided here, where the technical material is difficult enough as it is.

individuals and individuals, and so on. Where the type of an entity is obvious, we omit the superscripts.

One might worry that the present notions of property and relation are too extensional. The adequacy of this view in the present context should be judged on what inferences it licenses. Suppose, for example, that P and Q are predicates that are given properties as semantic values and that co-refer in all possible worlds. In this instance we would have to hold that, for example, the implication $P(i) \rightarrow A$ is equivalent to $Q(i) \rightarrow A$. This inference does not seem especially intuitive to me, but then again, I have no intuitions that weigh against it either. But if this inference is problematic, there are simple ways to complicate our theory of properties (for example, by making a property to be a pair of a property in the present sense and a Fregean mode of presentation or a mental representation, and so on). Until such a difficulty is found, I suggest that we resist all attempts to make the theory more complicated.

4.5 States of affairs

A state of affairs (SOA) is a fact-like entity. It is what makes statements true at situations. On my view a SOA is just a sequence of the form:

$$\langle P, e_1, \dots, e_n \rangle$$

where P is a relation of type $(\tau_1, \dots, \tau_n)$ and each e_i is an entity of the corresponding type τ_i . A sequence is just a set theoretic construct – it is an ordered set. Thus, along with the relation and entities that are its constituents, a SOA is reducible to the elements set out above.

Some philosophers, like David Armstrong, take SOA not to be reducible to entities of other categories (Armstrong 1997). Armstrong, however, has different work in mind for SOA in addition to that which I want them to do. I merely contend that *qua* set theoretic construct SOA will be adequate for the tasks at hand.

My SOA also differ from those standardly used in situation semantics. Most situation semanticists include a ‘polarity’ of 0 or 1 in their SOA. These polarities are like truth values. Thus, $\langle \text{Dog, ramsey}; 1 \rangle$ is the SOA that Ramsey is a dog and $\langle \text{Dog, ramsey}; 0 \rangle$ is the SOA that Ramsey is not a dog. The reason for including polarities is that they are used in the theory of negation that has been attached to situation semantics. In chapter 4 below, I present a theory of negation that does not require polarities. Hence I do not have any reason to include them in SOA.

SOA represent features or facts about worlds. They might do so accurately like the SOA, $\langle \text{Dog, ramsey} \rangle$ or they might pick out an impossibility such as $\langle =, \text{NaCl, water} \rangle$ and so do not accurately represent facts at any worlds. To be a bit more technical, we can define this representation relationship as

follows:

A SOA, $\langle P, e_1, \dots, e_n \rangle$ accurately represents a fact at a world w if and only if $(e_1, \dots, e_n) \in P(w)$.

Note that this definition does not assume that in addition to worlds there are atomic facts. It might be, à la Davidson, that worlds themselves are each one big fact. Subworld facts do not play any role in our theory and so I will not discuss them further.

In section 6.14 below, I will suggest that we might treat the constituents of SOA slightly differently than they are treated here. But, for the most part, we can assume the present simple theory of SOA for the remainder of the book.

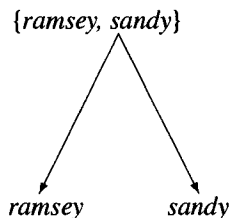
4.6 Non-well-founded set theory

The next element that we will need in our semantics is a theory of sets. We will use a non-standard theory, instead of a standard theory like Zermelo–Fraenkel set theory, Bernays-von Neumann–Gödel set theory, and so on. Unlike these set theories, the one that we will employ does not contain an axiom of foundation.

The axiom of foundation is an integral part of the *iterative* conception of a set.² According to the iterative conception, the set theoretic universe is divided into a strict hierarchy of levels. At the bottom level are the pure elements (Urelements), if the set theory admits any. At the second level, there are sets of Urelements. At successive levels, there are sets of elements taken from previous levels. In this way, no set that is a member of itself, or a member of a member of itself, etc., is ever admitted into any level of the hierarchy. The axiom of foundation, largely, enforces that no violation of the hierarchy restrictions takes place.³

Peter Aczel's non-well-founded set theory does not contain an axiom of foundation. Instead, it contains all the other axioms of Zermelo–Fraenkel set theory and the anti-foundation-axiom. We don't need to state these axioms here. Rather, I will merely present an intuitive picture of the theory, due to Aczel and used by Barwise and Etchemendy (Barwise and Etchemendy 1987).

We can think of a set in terms of tree diagrams, such as:

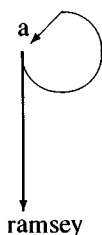


² My explanation of the iterative conception of a set is borrowed from (Fraenkel *et al.* 1973), pp 87ff.

³ But see (Parsons 1983). There are some complications that come with this interpretation.

which represents the set $\{\text{ramsey}, \text{sandy}\}$. A point on a diagram like this is called a 'node'. If there is a vertex going from one node to another the latter is called a 'child' of the former. Thus, for example, the node labelled with 'ramsey' is a child of that with the label ' $\{\text{ramsey}, \text{sandy}\}$ '.

The anti-foundation axiom says, roughly, that for any tree, there is a set that corresponds to it. The tree describing $\{\text{ramsey}, \text{sandy}\}$ is well founded. This means that there are no cycles in it. But not every tree is well founded. Consider, for instance, the following tree:



This diagram represents the set $a = \{\text{ramsey}, a\}$. This set is a member of itself. There is a tree describing it, so it is a set in the universe of Aczel's theory. There are many sets that have themselves as members, on this theory. There is even a set that has only itself as a member.

We will discuss my reasons for adopting non-well-founded set theory in the next section. Here we will outline some other uses to which this theory has been put.

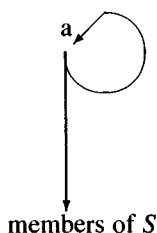
There are applications of non-well-founded set theory in mathematics, computer science and philosophy. Among philosophers, perhaps the best known application is the use of non-well-founded set theory by Barwise and John Etchemendy to construct a theory of truth (Barwise and Etchemendy 1987). But there are a lot of other circular phenomena that are of interest to philosophers upon which non-well-founded set theory can be brought to bear.

Here is one due to Barwise and Lawrence Moss (Barwise and Moss 1996). Consider the hypergame paradox (originally due to William Zwicker). A standard game is a game that always ends after a finite number of moves. Thus, for example, tick-tack-toe is a standard game. After every square is filled, it ends. Hypergame is played as follows. The first player chooses a standard game. Then the standard game is played. When the standard game is finished, so is hypergame. Thus it would seem that hypergame is a standard game, since it always ends with one move more than the standard game that is chosen in its first move. But, if it is a standard game, then it would seem that hypergame can be chosen in the first move of a hypergame. Then, if hypergame is chosen, a standard game must be chosen in move two. Suppose that hypergame is chosen in this move as well, and so on, ad infinitum. Thus, it would seem that hypergame does not always end in a finite number of moves, and so it is not a standard game. But

then, if hypergame is not a standard game, it cannot be chosen as the first move of a hypergame. Thus, it would seem that we again derive that hypergame is a standard game that takes one move longer than the standard game chosen in its first move. Therefore, hypergame is a standard game if and only if it is not a standard game.

This paradox is particularly annoying since it is easy to specify what a standard game is and what a hypergame is. It would seem that there is such a game. On the other hand, it seems to be both standard and non-standard. Barwise and Moss, however, think that there is an equivocation in the statement of the paradox. They think that there are two sorts of games that are being confused in the description of hypergame. Let S be a set of games. Then they define S^+ to be the *supergame* defined over S , i.e., S^+ is played by choosing one of the games in S and playing that game. A set is a well-behaved collection of entities. It obeys the axioms of set theory (in this case non-well-founded set theory). It can be shown that, if S is a set of standard games, then S^+ is not a member of S . On the other hand, let S^* be defined as the *hypergame* defined over S . That is, S^* is played by choosing a member of S or S^* itself and playing that game. Clearly, where S is a set of standard games, S^+ is a standard game, but S^* is not. What is learned is that the collection of standard games is not itself a set, and we cannot prove that there exist supergames or hypergames over it ((Barwise and Moss 1996), §12.3).

Non-well-founded set theory gives us an elegant way of modelling hypergames. Let S be a set of standard games. Then $S^* = S \cup \{S^*\}$. In a tree diagram S^* is pictured as



The hypergame defined by a set of standard games is not a paradoxical object. It is merely a game with an infinite number of moves. Of course we could model the infinite hypergame in standard set theory as an infinite sequence of repetitions of the word 'hypergame'. But the self-referential aspect of hypergames come out much more clearly in non-well-founded set theory.

As I said above, there are applications of non-well-founded set theory in mathematics and computer science. The mathematical applications are mainly in the very technical field of bisimulation theory, which I will not delve into here. But there are uses in computer science that are relatively easy to understand. For

example, non-well-founded set theory can be used to model *labelled transition systems*. Suppose that we want to model the evolution of a dynamic system (like a computer running a program). At any point in time, the system is in a state and an action is about to take place. Barwise and Moss formulate the theory of labelled transition systems by first taking a set of states, S , and a set of actions, Act . Suppose that s and t are states and a is an action. Let $\delta(s)$ be the set of transitions possible in a system from state s . If the pair (a, t) is in $\delta(s)$, then according to the system, it is possible to begin in state s and end up in t after the performance of a . We say then that a labels a transition from s to t (this is written $s \xrightarrow{a} t$). Barwise and Moss give the nice example of the system *clock*. Either one of two actions can take place in the starting position. Either the clock can tick and then it returns to the starting position or it can break, in which case it ends up in the state *stuckclock*. It makes sense to think of the first state as the state, *clock*. For then we know that if it returns to this state that we have the same possibilities of actions all over again. This sort of circularity is easily accommodated in non-well-founded set theory. We can think of this system as the set $clock = \{ \langle tick, clock \rangle, \langle break, stuckclock \rangle \}$ ((Barwise and Moss 1996), §3.2). Circularity of this form is ubiquitous in computer science, and so it is good to have a tool that deals with it in such a straightforward manner.

Thus, it would seem that there are independent reasons to accept non-well-founded set theory.

4.7 Constructing abstract situations

We put non-well-founded set theory to work in constructing the abstract situations for our model theory. Our construction follows Barwise and Moss in their use of non-well-founded set theory to construct models for modal logics.

A situation is an ordered pair $s = (SOA(s), R(s))$. $SOA(s)$ is a set of states of affairs; these are the states of affairs that obtain at s . $R(s)$ is a set of ordered pairs. (t, u) is in $R(s)$ if and only if $Rstu$ holds in our frame. Thus, the set of situations itself determines which situations are related to which in the frame. We need non-well-founded set theory for this construction because for each situation s , $Rsss$. This means that (s, s) is in $R(s)$ and so, s is not a well-founded set.

We can now define the informational part-whole relation on situations discussed in chapters 2 and 3 above. We say that $s \sqsubseteq t$ if and only if $SOA(s) \subseteq SOA(t)$ and $R(t) \subseteq R(s)$. The first half of this definition seems obvious. If a situation s contains less information than t , then all of the states of affairs that hold in s also hold in t . The second half might seem less intuitive, but it says something similar. The fewer pairs of situations related to a given situation u ,

the more implicational propositions hold at u . Thus, if $R(t)$ is a subset of $R(s)$ then at least as many implications hold at t as hold at s .

4.8 Possible situations

We also need the notion of a possible situation, both for what we have already said in chapters 2 and 3 above and for our theory of negation set out in chapter 4 below.

Before we can define a possible situation, we will need to understand what it means to say that a situation minimally obtains at a world:

Definition 4 (minimally obtains) *A situation s minimally obtains at a world w if and only if (a) for all σ in $SOA(s)$, σ holds at w and (b) for all situations t , if $SOA(s) = SOA(t)$, then $s \trianglelefteq t$.*

In other words, a situation minimally obtains at a world if it represents that world in terms of its SOA and, if from the set of situations that have the same SOA as it does, it makes the fewest implications true. The reason we need this definition is that there are cases in which all of the SOA in a situation hold at a world, but the situation itself is not possible.

Now, we can define a possible situation.

Definition 5 (Possible Situation) *A situation s is possible if and only if there is some situation t such that there is a world at which t minimally obtains and $s \trianglelefteq t$.*

The idea here is quite simple. A possible situation is just a situation that obtains at some possible world, or perhaps in more than one possible world.

I also suggest that there is a strong relationship between possible worlds and logical situations. This relationship is summed up in the following claim:

Claim 6 *For each world w there is a logical situation s such that there is some situation t for which both $s \trianglelefteq t$ and t minimally obtains at w .*

Or, in simpler language, for each world there is a logical situation that obtains at it. I am not exactly sure how to justify this. Some support comes from its consequence that all valid formulae are also necessarily true.

Perhaps we could rely on our theory of situated inference for help in justifying our claim. It seems reasonable to suggest that in each possible world there are the salient facts about the intended model as a whole. In our world there are the truths about the set-theoretic universe (if there is such a thing) and these facts are in every possible world. Moreover, in every possible world there are facts about the set of situations and the structures that we are constructing here. Thus,

in each world there is information about what propositions semantically entail which propositions in the model. In each possible world, therefore, the links are available to make true ' $A \rightarrow B$ ' where ' A ' semantically entails ' B ' in the intended model.

4.9 The metaphysics of informational links

In chapter 2 above, I suggested that situated inference utilises informational links. Among those links are laws of nature and conventions. In this section, we will talk about the way in which informational links enter into abstract situations.

The starting point for my view on the metaphysics of informational links is the theory of laws of nature due to David Armstrong, Fred Dretske and Michael Tooley: see (Armstrong 1983), (Dretske 1977), (Tooley 1977) and (Tooley 1987). On the Armstrong–Dretske–Tooley view of laws, a simple law of nature has the form,

$$N(P, Q)$$

where P and Q are properties and N is a second-order property, that of nomic necessitation. ' $N(P, Q)$ ' is to be read 'Being- P nomically necessitates being- Q '. On the Armstrong–Tooley–Dretske view, laws of nature are properties being related together. Nomic necessitation, in this view, is a primitive, not to be analysed in terms of anything more basic.

I adopt from this theory the view of the 'logical form' of laws of nature, but not necessarily its view of the ontology of laws. I hold that when we think that a law holds, we think that it holds not just in a particular instance between particular things, but it holds because of the properties of those things. Thus, our understanding of laws of nature are in terms of relations between properties. Whether in reality these relations have any more basic origin or analysis I do not know, and for our semantic purposes it does not matter.

Let us examine a simple example to see how this view can be incorporated into our theory. Suppose that the law-like SOA, $\langle N, P, Q \rangle$, is in a situation, s . Also, suppose that i is an individual and that P and Q are properties of individuals. Then the implication ' $P(i) \rightarrow Q(i)$ ' is true at s .

Note that in our ontology we need many different nomic necessitation relations, for not all laws of nature or possible laws of nature can be formulated as a relation's holding between two monadic properties of individuals. Consider, for example, the Newtonian 'law' that all matter attracts all other matter. One way that this might be formulated as a SOA is $\langle N, \rangle (d(x, y), 0), \textit{Attracts} \rangle$. Here, *Attracts* and $\rangle (d(x, y), 0)$ (the distance between x and y) are binary

relations.⁴ Thus, N is a binary relation between binary relations. Therefore, in order to provide a satisfactory theory of laws of nature, we will have to commit ourselves to a very wide range of types of nomic necessity relations.

This commitment might be a problem for Armstrong's theory, but not for mine. On Armstrong's view, the necessitation relation stands between universals and is itself a universal. On Armstrong's view there are not as many universals as there are properties and relations on my view. He has to be careful about postulating new relations. I do not have to be very careful at all about postulating relations. Functions are cheap, and relations are just functions of a certain sort.

With regard to conventional links my view is similar. There is a family of relations *Con* that hold between properties and other relations (or perhaps between individuals, events, and so on). Some SOA that represent conventional links will be very complicated, but as I said above, relations are cheap and we can construct whatever is necessary to provide adequate formulations.

Having an informational link in a situation constrains both the other SOA in the situation and the way in which that situation is related to other situations. Links provide information about what implications are true at a situation, and so they have to constrain what situations are related to which. For example, suppose that $\langle N, P, Q \rangle$ is in s . Then, if $Rstu$ and $P(i)$ holds at t , ' $Q(i)$ ' must hold at u . Moreover, if ' $P(i)$ ' holds at s itself, ' $Q(i)$ ' must obtain there as well.

Exactly which metaphysics of informational links is adopted makes little difference for my theory as long as it is very intensional. It cannot hold that the informational links that obtain at a situation are closed under classical logic or even under a standard form of strict implication. Armstrong, Tooley, and Dretske's theory of laws of nature does not have this consequence and that is why I adopt it. But any other theory that also treats these links as very intensional entities will do as well.

4.10 Ersatzism

In the terminology of Lewis (Lewis 1986), our situations are *ersatz*. Lewis distinguishes between vertebrate realist theories and ersatzist theories about possible worlds. Vertebrate realist theories hold that other possible worlds are places like our own universe. They contain individuals standing in relation in space–time to one another. Ersatzist theories, on the other hand, claim that other possible worlds are either abstract objects or abstract constructions, such as sets of sentences or sets of SOA. Clearly, our situations are ersatz.

⁴ More properly, perhaps we should have $\lambda x y (> (d(x, y), 0))$, but this doesn't really matter here. What might matter, however, is that this expression might stand for a construct of properties and functions, not a primitive property and this would make the SOA even more complicated than I indicate in the text.

David Lewis argues that there are serious problems with ersatzism about possible worlds (Lewis 1986). Luckily for us, these problems do not affect our theory of abstract situations. There are three major criticisms that Lewis directs at ersatzism. Lewis does not direct all these criticisms at every ersatzist theory, but let us see how our view avoids each of them.

The first criticism is that ersatz theories do not explain how ersatz worlds represent real worlds (ibid., ch. 3). I, however, have presented a theory of how situations represent worlds (see section 4.8 above).

The second criticism is that ersatz theories take modality to be a primitive (p. 176). On our theory, there are two ways to interpret this criticism, for there are two basic sorts of modality on our view. The first type we might call *absolute modality*. A situation is absolutely possible if and only if it is in our set of possible situations. Since we take possible worlds to be primitives, and these determine which situations are possible, we are in as good a position to explain absolute possibility as any theory of possible worlds. Any theory of possible worlds can be grafted onto the present theory, and it can be used to explain modality. There is a second sort of possibility that we will discuss in chapter 6 below. This is *relative modality*. A situation may contain the information that another situation (or itself) is possible in some sense. We will elaborate on the sense in which this can occur in chapter 6.

The third criticism is that ersatz theories do not explain how their worlds can represent possibilities. For example, suppose we have a view that metaphysical possibilities are determined by the different combinations of basic physical properties. On this view, how are we to tell whether an ersatz world represents there being a talking donkey (ibid., p. 188)? Luckily, we have an easy answer to that question too. On our theory, the predicate 'is a donkey' represents a property in our semantics. This property is just the function from worlds to the set of singletons of donkeys in those worlds. It would seem reasonable to hold that 'is a donkey' represents a property, since as Kripke, Putnam, and others have argued, species terms are *directly referential*. Directly referential terms refer to things or properties without having any descriptive content; they merely refer. If 'is a donkey' merely refers, it makes sense to say that it merely refers to the property of being a donkey. Let D be this property. Thus, whether there is a donkey in a particular situation, s , in our ontology is determined by whether there is some individual i such that $\langle D, i \rangle$ is in s . The semantics of the predicate 'talks' is, no doubt, more complex. Perhaps 'talks' represents a set of interrelated properties like 'makes noise with its mouth', 'this noise is intelligible' and so on, or perhaps 'talks' represents a Wittgensteinian family of such properties. At any rate, I suggest that all predicates can be given a satisfactory semantics in terms of the properties in my ontology. If this suggestion is correct, then we only need look at a situation and see whether those properties needed to interpret 'talks' are had by i in s .

With regard to impossible situations, Lewis compares the claim that there are worlds in which contradictions come true to there being a distant mountain on which there are true contradictions.

For comparison, suppose travellers told of a place in this world – a marvellous mountain, far away in the bush – where contradictions are true. Allegedly, we have truths of the form ‘On the mountain both P and not P ’. But if ‘on the mountain’ is a restricting modifier, which works by limiting domains of implicit and explicit quantification to a certain part of all that there is, then it has no effect on the truth-functional connectives. Then the order of modifier and connectives makes no difference. So ‘On the mountain both P and Q ’ is equivalent to ‘On the mountain P , and on the mountain Q ’; likewise ‘On the mountain not P ’ is equivalent to ‘Not: on the mountain P ’; putting these together, the alleged truth ‘On the mountain both P and not P ’ is equivalent to the overt contradiction ‘On the mountain P , and not: on the mountain P ’. That is there is no difference between a contradiction within the scope of a modifier and a plain contradiction that has the modifier within it. So to tell the alleged truth about the marvellously contradictory things that happen on the mountain is no different from contradicting yourself. But there is no subject matter, however marvellous, about which you can tell the truth by contradicting yourself. ((Lewis 1986), p. 7n)

This comparison between an impossible world and a distant contradictory mountain holds only if we think of worlds like other bits of space–time. If we think of them as collections of sentences, propositions or SOA, then the problem does not arise. For example, we do not contradict ourselves if we say that there is a set that includes a statement and its negation. Thus ersatzers, unlike Lewis, can help themselves to impossible worlds or impossible situations. We will look at Lewis’ argument again in greater depth in chapter 5 below.

4.11 Propositions

A proposition is a piece of information of a certain sort. A proposition is a set of situations closed upward under $\leq$. This means that if φ is a proposition and s in φ , then for any t such that $s \leq t$, t is in φ as well.

This notion of a proposition might seem somewhat strange, when coupled with the fact that situations were introduced to deal with semantic partiality. Consider an example that we discussed in chapter 3 above. Suppose that I say

The dog barked at a possum yesterday.

Appealing to the situation in which I make this statement will allow us to determine the reference of ‘the dog’. It does not refer to the only dog in the universe, but to my dog, who is one dog among many (although an extraordinary dog, of course). The word ‘yesterday’, when said in a situation that takes place today refers to the eleventh of February 2003. If we expand the situation in

which I made the statement to include more information, then ‘the dog’ may fail to refer. To create this failure, we merely have to include information about other dogs. Then the situation will no longer contain a single dog. Similarly, if we enlarge the situation to include days other than today the reference of ‘yesterday’ will no longer be clear.

To understand partiality, we need to distinguish between the information conveyed by a statement and the form of the statement itself. We first locate the statement in a context of utterance to determine the referents of the denoting phrases that it contains. Then we can determine what proposition is conveyed by the statement.

One might object that this shows that we need situations only as contexts of utterance. They are not needed to represent the contents conveyed by those utterances. We could merely take situations to be parts of worlds and take sets of worlds to be propositions.

To take this position is to ignore that we convey partial information about the world to one another all the time. If we take propositions to be sets of worlds, then every proposition ‘includes’ (in the sense of entailing) everything that is true at all possible worlds. In every statement we make, on the possible worlds view, we also assert the truth of every necessary truth. The examples with which we began this book show that this view is counterintuitive. One nice thing about situations, from a logical point of view, is that they allow us to make much finer discriminations between truths. Since we already allow situations into our theory as contexts of utterance, it seems to do little harm and much good to allow them to act as constituents of propositions as well.

4.12 Doing without non-well-founded set theory

We can reconstruct the theory using standard set theory, rather than non-well-founded set theory. I use non-well-founded set theory because the resulting view is mathematically more elegant. As I said in chapter 3 above, I think that when we are choosing our theories, we should take elegance into account.

I developed a version of the present theory based on standard set theory in (Mares 1996). In that paper, I characterised isolated inference using the relation *Involves* between types of ‘infons’. The notion of an infon is due to Keith Devlin. An *infon* is a piece of information. I treated situations as sets of infons. We can define our ternary accessibility relation using infons. Suppose that s , t , and u are situations. $Rstu$ if and only if, for infon types T and T' , if the infon (*Involves*, T , T') is in s , if there is an infon σ of type T in t , then there is an infon σ' of type T' in u .

Informational closure conditions were then imposed on situations themselves. For example, for all situations s example, if (*Involves*, T , T') is in s

and so is $(\text{Involves}, T', T'')$, then $(\text{Involves}, T, T'')$ is also in s . On this theory, the closure conditions on situations do all the work and the ternary relation is quite redundant but harmless (for what is a ternary relation but a set of ordered triples?). I prefer the cleaner theory that uses non-well-founded set theory. For those, however, who are horrified by the idea of giving up the axiom of foundation, there is this alternative interpretation of relevant logic.

5 Negation

5.1 Impossible situations and incomplete situations

As we have said several times, one of the central motivations behind the development of relevant logic is to avoid the paradoxes of implication. Consider two of these:

$$(S2) A \rightarrow (B \vee \sim B)$$

$$(S3) (A \wedge \sim A) \rightarrow B$$

Given our semantics for implication and disjunction, the only way that we can avoid making (S2) valid is to accept situations which are not bivalent. That is, we need situations at which, for at least some formula B , neither B nor its negation is true. Another term for a situation which is not bivalent is a ‘partial situation’. Similarly, in order to reject (S3) we must accept situations which are inconsistent. We need situations which *do* make true $A \wedge \sim A$. In other words, we need a semantics for negation that allows there to be partial situations and inconsistent situations.

But consider the standard semantics for negation. In classical logic with its usual semantics, the truth condition for negation treats negation as ‘failure’, i.e.,

‘ $\sim A$ ’ is true in s if and only if ‘ A ’ fails to be true in s .

On the face of it, this truth condition would force us to make every situation bivalent and consistent. For, each situation either makes A true or fails to make it true, and no situation does both.

Graham Priest has argued, however, that this truth condition does not force us to accept the bivalence and, in particular, the consistency of every situation (Priest 1990). He points out that these consequences of the truth condition depend crucially on the metalanguage used. If we take the word ‘fails’ in the standard truth condition itself to be a non-classical negation (such as the one that we are going to develop in this chapter), then we need not accept consistency or bivalence. But that is not the approach that we are taking in this book. I am

using a completely classical metalanguage. I will justify this use of classical logic in chapter 10 below. But for now I will repeat the motivation that I used in chapter 1 above. Classical logic is the most widely understood and most thoroughly studied logic. Its use in examining non-classical logics makes the non-classical logics more intelligible, both for the beginner and the seasoned non-classical logician.

The issue of a metalanguage having been set aside, we now have to find a semantics for negation to replace the standard classical semantics. Our treatment of this issue will not follow the chronological development of the treatment of negation in relevant logic. Rather, we will discuss my favourite approach to negation first and then later examine how this approach fits into the history of the subject.

5.2 Dunn's compatibility semantics

Dunn treats negation in terms of a binary relation between situation. This relation, *Cst*, says that the situations *s* and *t* are compatible with one another. Dunn's compatibility relation has its roots in Rob Goldblatt's semantics for quantum logic¹ as well as Richard and Val Routley's semantics for negation for relevant logic. We will discuss the Routleys' view at length in the section 5.4. Let us begin here with Goldblatt's semantics, since Goldblatt's idea is very close to the intuitions that I think lie behind negation in relevant logic.

Quantum logic originates with mathematics of quantum physics. This sounds rather complicated, and is, but we need only understand some very basic notions for our purposes. Goldblatt's semantics is a frame theory, like Kripke's semantics for modal logic or the Routley–Meyer semantics for relevant logic. That is, it contains a set of indices and some relations on those indices. Here the indices are to be understood as the possible outcomes of experiments of the sort done by quantum physicists. Some of these outcomes are compatible with one another and some are not. We do not need to examine examples of compatible and incompatible outcomes of experiments, since the sorts of cases we are interested in are, by and large, much less complicated than the examples from physics. And we will examine some such simpler examples soon.

Goldblatt adds to the set of outcomes a binary relation, $\perp$, which is called 'orthogonality'. This relation captures the notion of two outcomes' precluding one another. In other words, where *a* and *b* are possible outcomes of experiments, $a \perp b$ means that *a* and *b* preclude one another from occurring. That is,

¹ Goldblatt actually produces a semantics for *orthologic*, which is a generalisation of quantum logic.

if a occurs then b will not occur. We then have the following truth condition for negation:

‘ $\sim A$ ’ iff is true in s if and only if $\forall x(A \text{ is true in } x \supset a \perp x)$.

In other words, a makes ‘ $\sim A$ ’ true if and only if all outcomes which make true ‘ A ’ are precluded by a ((Goldblatt 1974), p. 86).

Moving from quantum logic to relevant logic, we can still think of a situation as precluding other situations. One situation can preclude another situation from obtaining at the same world. For example, suppose a situation s says that my dog Ramsey is black all over (at a particular time) and a situation t says that he is white all over (at the same time). Then s and t are incompatible, since being white all over is incompatible with being black all over. Here we have a notion of incompatibility that is a little more common-or-garden than that used to motivate the semantics for quantum logic, but the idea is the same; some situations are compatible with each other and others are not.

Instead of treating negation in terms of incompatibility, as Goldblatt does, Dunn looks at its complement, *compatibility* (C). Thus, Cst holds if and only if s and t are not incompatible. So, we now have the following truth condition for negation:

$s \models \sim A$ iff $\forall x(Csx \supset x \not\models A)$.

A situation makes $\sim A$ true if and only if every situation compatible with it fails to make A true.

How does the use of compatibility help with our problems of finding ways to make situations partial and inconsistent?

Partiality is straightforward, so we will start with that. Consider the situation that consists of the information that is currently available to me. This includes what is going on in my study as I write this section of my book, and what I can see through my window. Nothing happening here makes it true that it is currently raining in Toronto (which is on the other side of the globe). But situations in which it is raining in Toronto are compatible with my current situation. So neither ‘It is raining in Toronto’ nor ‘It is not raining in Toronto’ is true in my current situation. Thus, bivalence fails for this situation and the situation is partial.

Inconsistency is a little stranger. Consider a situation s in which my dog Ramsey is both black and white all over at the same time. The way in which we constructed situations in chapter 4 above allow this sort of situation. The situation s is not compatible with itself! Put like this it sounds bizarre, but consider someone who says incompatible things about something, say, about the colour of Ramsey’s coat. This person’s representation of the world is incompatible with itself. Our situations are representations and they can be incompatible with themselves in very similar ways – they can represent worlds in incompatible

ways. When a situation s is incompatible with itself, it is possible for it to make a formula A true, but A fail to be true in every situation compatible with s . In such cases, according to our truth condition for negation, s makes both A and its negation true. Thus, s is inconsistent.

Let's look at an example of a negated sentence and its truth condition. The following was suggested by an anonymous referee:

Oswald did not act alone.

On our view, if this sentence is true, it is so because something happened in an actual situation that precludes situations in which Oswald did act alone. For example, if there was a second gunman then there is an actual situation that contains both Oswald and the second gunman shooting at Kennedy. This situation is incompatible with any situation in which Oswald acted alone, and hence the above sentence is true.

Note that situations in which Oswald acted alone are not merely situations in which the only information on the topic that we have is that Oswald shot Kennedy. These are situations in which Kennedy was shot by Oswald, but do not preclude another gunman. For a situation to make true the statement 'Oswald acted alone', it must include information that precludes there having been any other gunmen, gunwomen, helpers and so on.

5.3 Compatibility and negation

This approach to negation derives in part from the doctrine of Plato's *Sophist*. At *Sophist* 256–9, Plato argues that to say that something is not such-and-such is to assert that it is somehow different from being such and such. Plato has one of the characters from his dialogue, the Stranger, say

when it is asserted that a negative signifies a contrary, we shall not agree, but admit no more than this – that the prefix 'not' indicates something different from the words that follow, or rather from the things designated by the words pronounced after the negative. (257b–c)

For example, when we say that something is not taller than something else we mean that the thing is short or medium sized compared to that other thing (257b).² Now, by 'difference' here Plato must mean a sort of incompatibility. Being red is different from being round, but to say that something is not round means that it is some other shape or has no shape, for these are incompatible with being round.

A similar theory is put forward more recently by Huw Price (Price 1990). He claims that we have negation to indicate incompatibility. For example, suppose that someone asks whether Fred is in the kitchen and is told in reply

² I am adopting Laurence Horn's interpretation of *Sophist* 257b here. See (Horn 1989), p. 5.

that Fred is in the garden. Knowing that the kitchen and garden are distinct places, the questioner knows that Fred is not in the garden. For being in the garden is incompatible with being in the kitchen. Price claims that our use of negation evolved to express this sort of incompatibility. Price's view is attractive because it presents a plausible psychological and social function for negation.³

There are important differences, however, between Dunn's view, Plato's view, and Price's view. All treat negation in terms of incompatibility, but on Dunn's view the compatibility holds between situations, on Plato's it holds between properties, and on Price's it holds between propositions. If all the sorts of incompatibilities needed for the semantics of negation can be reduced to incompatibilities between properties, then we can reduce Dunn's view to the Platonic view. This would be nice, but I am not sure how this reduction would go. Complex examples, such as the Oswald example discussed in the previous section, defy easy reduction to incompatibilities between properties, unless we allow complex properties such as 'acting in concert with another person' as relata of incompatibilities.

With regard to Price's view, propositions in the sense that we have them in our theory, are not the sorts of entities that can provide the basis of a theory of negation. Propositions are merely sets of situations. If we claim that two propositions are incompatible, then we need to say what makes them incompatible.⁴ Clearly, this must have something to do with the situations that they contain. But we can cash out the incompatibility between propositions in terms of an incompatibility between situations. Formally, we get the following relationship. *For any situation s , if s is in the proposition expressed by A , then for any situation t in the proposition expressed by $\sim A$, s and t are incompatible.*

One might object that our view merely takes the incompatibility between situations as primitive and so provides no more explanation of negation than does Price's. In a sense this is correct. I do take incompatibility to be a primitive relation between situations. Explanations have to come to an end somewhere, and so we have to have some unexplained entities in our theory – our 'primitives'. Our choice of primitives, however, can be good or bad. Some primitive notions are more intuitive than others. Let's return to the property-incompatibility theory for a moment. It seems quite reasonable to say that certain properties are

³ In fact my agreement with Price is only partial. He thinks that negation also expresses denial in addition to incompatibility. This I think is incorrect. As we shall see below, when we deny a statement, we are claiming that it does not hold. Saying that it is incompatible with the present situation makes a stronger claim.

⁴ In one sense, we do have an answer to this question. What makes them incompatible is that they do not contain any possible situations in common. But this answer, although clear and extensional, is not very illuminating in a philosophical sense.

incompatible with one another and this incompatibility is essential to those properties. Being black and being white are essentially incompatible with one another. Being square and being round are also incompatible with one another. I can't explain what this incompatibility is in any more basic terms, but I think it is clear what sort of thing I am describing. Likewise, incompatibilities between situations are intuitively clear, even if we cannot reduce these incompatibilities to incompatibility relations between pairs of properties.

5.4 Constraints on compatibility

The compatibility relation is a relation on situations. Like the accessibility relation, we can ask what properties the compatibility relation should have.

We have already ruled out the idea that it should be reflexive. That is, there are some situations that are not compatible with themselves. But it should be symmetrical, that is, for all situations s and t ,

If Cst , then Cts .

We have treated ' Cst ' as saying that s and t are compatible with each other, and being compatible with each other is clearly symmetrical. The symmetry of C makes valid the inference at a situation from any proposition A to its double negation $\sim\sim A$. In terms of our natural deduction system, it makes valid the rule of double negation introduction, viz.,

($\sim\sim I$) From A_α to infer $\sim\sim A_\alpha$.

The other direction of double negation, viz., the inference rule,

($\sim\sim E$) From $\sim\sim A_\alpha$ to infer A_α ,

is much more difficult to make valid. The semantic postulate that corresponds to the rule of *double negation elimination* is, for all situations s there is some situation x such that Csx and for all situations y such that Cxy , $y \leq s$ ((Mares 1995), p. 587). This postulate is complicated and rather inelegant.

There is, however, an easier way of capturing double negation elimination. We can add a postulate that for each situation s there is a maximal situation t such that Cst . More formally, we can add the postulate,

(star postulate) $\exists x(Csx \ \& \ \forall y(Csy \supset y \leq x))$.

The reason for the name of this postulate will become obvious soon.

Given the star postulate, we can define an operator on the set of situations. An operator is a mathematical entity which takes an element from a set and returns a unique element from the same set. The operator that we define here

is the ‘Routley star operator’, so-called because it was introduced by Richard and Val Routley in (Routley and Routley 1972). The star of a situation s , s^* , is the maximal situation compatible with s . The star postulate ensures that there is an s^* for every situation s . If we add the star postulate and star operator to our semantics, we can derive the following equivalence:

$$‘\sim A’ \text{ is true in } s \text{ iff } ‘A’ \text{ fails to be true in } s^*.$$

And we can add the following postulate to make valid double negation elimination:

$$s^{**} = s.$$

Clearly, the above postulate (which, in algebraic terms says that the star operator is of ‘period two’) is much more intuitive than the semantic postulate stated merely in terms of the compatibility relation. The star operator has great formal virtues. It enables us to produce an elegant semantics for a DeMorgan negation. A negation is said to be DeMorgan if it has double negation introduction and elimination and satisfies the usual DeMorgan laws, such as $\sim(A \wedge B) \leftrightarrow (\sim A \vee \sim B)$ and $\sim(A \vee B) \leftrightarrow (\sim A \wedge \sim B)$.

The star operator has been the subject of much criticism. The criticism has been philosophical. Jack Copeland and Johann van Benthem have complained that the star operator has no intuitive meaning, and so it does not provide a philosophical interpretation of negation in relevant logic (Copeland 1979), (van Benthem 1979). Dunn replies to Copeland and van Benthem by giving an intuitive interpretation of the star operator. The star operator takes a situation to the maximal situation that is compatible with it.

This, however, does not end the issue. The introduction of the star postulate naturally raises the following question. Should we believe that, for every situation s , there is some maximal situation s^* that is compatible with it? I am not sure how to argue for either an affirmative or a negative answer to this question. Perhaps the intuitiveness of double negation elimination supports the acceptance of the star postulate. On the other hand, perhaps relevant logicians should take seriously the idea of abandoning double negation elimination and some of the DeMorgan laws and instead adopting some intuitionist-like negation.

5.5 Worldly situations

A worldly situation is a logical situation that *exactly covers* a world. That is, a logical situation s is worldly if and only if there is some world w such that s minimally obtains at w and, for all situations t that obtain at w , $t \trianglelefteq s$.

We postulate that the worldly situations are those situations that are *maximally consistent*. They are consistent, so they are compatible with themselves, but they are also complete. That is, if s is worldly and Cst , then t is part of s (i.e. $t \sqsubseteq s$). Let's justify this condition. Suppose that s is worldly and Cst . That is, t is compatible with s . Assume, for the sake of a reductio, that t is not part of s . Then t contains some information that is not in s . Thus, there is information that is not in s but does not conflict with the information in s . But this means that s is not complete. And so, if worldly situations are to be considered complete, then they are only compatible with situations that are their parts.

When we add the definition of the star operator and the postulates given above, we find something interesting happens. For all worldly situations, s , we can show that,

$$s = s^*.$$

Thus, given the truth condition for negation phrased in terms of the star, we can show that, for all worldly situations, s ,

$$' \sim A ' \text{ is true in } s \text{ iff } ' A ' \text{ fails to be true in } s.$$

This is the standard, classical truth condition for negation. From this fact, we can show that, if s is a worldly situation, it satisfies two important semantic principles. The first of these is the principle of bivalence, viz.,

$$\text{Either } ' A ' \text{ or } ' \sim A ' \text{ is true in } s.$$

The other is the principle of consistency:

$$\text{It is not the case that both } ' A ' \text{ and } ' \sim A ' \text{ are true in } s.$$

Thus, when we restrict ourselves to worldly situations, we find that negation behaves classically.

At worldly situations disjunction and conjunction also behave classically.⁵ We then have all instances of the following schemes true at all worldly situations: the law of excluded middle, viz.,

$$A \vee \sim A,$$

and the law of non-contradiction, i.e.,

$$\sim (A \wedge \sim A).$$

⁵ It might seem that both of these connectives behave classically at all situations. Jack Copeland has argued that relevant disjunction can only be said to behave intuitionistically at partial situations (see (Copeland 1983)). I don't think this point is very important, so I will give it to him. Clearly, however, at worldly situations, classical disjunction and relevant (extensional) disjunction behave the same way.

In fact, all the theorems of classical logic that can be expressed in the fragment of our logical language that includes only propositional variables, conjunction, disjunction, negation, and parentheses are true at all worldly situations.

In fact, it can be shown that all theorems of classical logic (formulated in this vocabulary) are true at all logical worlds. In proof-theoretic terms, by the completeness theorem, it can be shown that all theorems of classical logic are theorems of the logic *R* of relevant implication – the logic that we are examining in this book. This restricted vocabulary is all that is needed to express all classical tautologies, in the sense that we can express any two-valued truth function using this vocabulary. Thus, we can think of our relevant logic as an extension of classical logic – it includes all the theorems of classical logic plus some other theorems in the vocabulary that includes implication. Of course, if we think of classical logic as the system that is formulated using this vocabulary and the arrow as material implication, we can also think of classical logic as an extension of our relevant logic.

In addition, the classical behaviour of these connectives allows us, when restricting inferences to being about worldly situations, to use all the classical rules of deduction. We can use, for example, disjunctive syllogism in its usual form. That is, for any worldly situation *s*, if '*A* ∨ *B*' and '∼*A*' are both true in *s*, then so is '*B*'. In fact, when we restrict our language to just what can be expressed with propositional variables, conjunction, disjunction and negation, both the set of formulae true at worldly situations and the set of rules under which these formulae are closed are exactly those of classical logic.

5.6 *A formal reduction of incompatibility

(This section concerns a formal 'nicety'. Nothing else in the book depends upon it.)

Whereas we take incompatibility to be philosophically basic, we do not have to have Dunn's compatibility relation as a primitive in our formal semantics. In (Mares 1995) I show that we can treat negation in the semantics for the logic *R* taking as our semantic primitives only the set of situations, the set of worldly situations, and the ternary accessibility relation, as well as a value assignment. Negation can be treated in terms of those primitives alone. The idea is fairly simple. Consider the class of possible situations. All the situations that are not in this class – the impossible situations – can be treated as a proposition. Let's call this proposition, *f*. We can read *f* as saying 'Something impossible occurs', since something impossible happens at all and only impossible situations.

Using the constant *f* we can define negation:

$$\sim A =_{df} A \rightarrow f.$$

This sort of implicational definition of negation will be familiar to some readers from the literature on intuitionist logic (and from some treatments of classical logic). But unlike the semantics for intuitionist and classical logics, our false constant (f) expresses a proposition that actually contains situations.

Given our definition of negation, we can derive the following truth condition

$$\begin{aligned} & \text{'}\sim A\text{' is true in } s \text{ if and only if } \forall x \forall y ((R s x y \ \& \ \text{'}A\text{' is true in } x) \supset \\ & \text{'}f\text{' is true in } y). \end{aligned}$$

In words, ' $\sim A$ ' holds at s if and only if s contains the information that from the hypothesis that A is true in a situation in the same world an impossible situation obtains in that world.

We now can define an incompatibility relation. We read ' Nst ' as saying that s and t are incompatible, and define N as follows:⁶

$$Nst \text{ if and only if } \forall x (Rstx \supset \text{'}f\text{' is true in } x).$$

And we can prove that

$$\text{'}\sim A\text{' is true in } s \text{ if and only if } \forall x (\text{'}A\text{' is true in } x \supset Nsx).$$

Two situations s and t are compatible (Cst) if they are not incompatible. Thus, we can also prove that

$$\text{'}\sim A\text{' is true in } s \text{ if and only if } \forall x (Csx \supset \text{'}A\text{' fails to be true in } x),$$

and this is just Dunn's truth condition for negation.

It is nice that we can prove that negation in our relevant logic acts in somewhat the same way as negation in intuitionist and classical logics – that it can be treated in terms of implication and a false constant. Of course there are differences too. As we have said, our false constant ('something impossible occurs' or ' f ') may express a non-empty proposition. And the implication that is involved here is relevant, rather than material or intuitionist implication. But it does mean that we can sometimes appeal to similar intuitions to support our view of negation as classical logicians and intuitionists do to support theirs.

5.7 Not all contradictions are created equal

There are many other views of negation that hold that all impossibilities are equivalent to one another, that they all express the same proposition. This is a view that we deny. The argument for this view is that all contradictions

⁶ To do this we also need the assumption that for all situations s and t , there is some situation u such that $Rstu$.

entail every proposition, thus they entail one another. What we deny is that all contradictions entail every proposition.

There is an argument for the view that contradictions entail everything that has been attributed in the literature to both Bertrand Russell and the mathematician G. H. Hardy. Here is a version of the story due to Harold Jeffreys. The neo-Hegelian philosopher, J. M. E. McTaggart, is supposed to have asked Hardy to derive that he and the Pope are the same from $2 + 2 = 5$. Jeffreys presents the story:

McTaggart is said to have denied [that every proposition is entailed by a contradiction], saying ‘If twice 2 is 5, how can you prove that I am the Pope?’ G. H. Hardy answered ‘If twice 2 is 4, $4 = 5$. Subtract 3; then $1 = 2$. But McTaggart and the Pope are two; therefore McTaggart and the Pope are one.’ ((Jeffreys 1973), p. 18)

Although Hardy doesn’t here pretend to show that every contradiction is equivalent to every other contradiction, this argument seems to come uncomfortably close to this result. It can be used to show that any false arithmetical equivalence implies any identity statement. This is an intolerable irrelevance.

Let’s formalise Hardy’s argument in our natural deduction system. We use the notation ‘ $Cd(X)$ ’ to mean ‘the size of the set X ’ (‘ Cd ’ represents the operation *the cardinality of*). So we have

- | | |
|---|------------------------------|
| 1. $2 \times 2 = 5_{\{1\}}$ | <i>hyp</i> |
| 2. $1 = 2_{\{1\}}$ | subtracting 3 from each side |
| 3. $Cd(\{a, b\}) = 2_{\{2\}}$ | <i>hyp</i> |
| 4. $Cd(\{a, b\}) = 1_{\{2,3\}}$ | 2, 3, transitivity of = |
| 5. $a = b_{\{2,3\}}$ | 4, (?) |
| 6. $2 \times 2 = 5 \rightarrow a = b_{\{1\}}$ | $2 - 5, \rightarrow I$ |

Clearly, in order to make this argument convincing we would have to devise a rule that allows the derivation of step 5. To do so, we would have to discuss the logical principles governing the relationship between the size of a set and the identity or non-identity of its members. But we won’t do that here. There is another problem with this argument that is of greater interest to us.

The rule of transitivity of identity that we appeal to in the argument is the following:

From $i = j_\alpha$ and $k = j_\beta$ to infer $k = i_{\alpha \cup \beta}$.

A relevant logician, however, should reject this version of transitivity. For consider the following argument. Let r be Ramsey, a be ‘the most active dog in the world’, and l be ‘the laziest dog in the world’. If we accept the above version

of transitivity, we would have also to accept the following:

1. $a = r_{\{1\}}$ *hyp*
2. $l = r_{\{2\}}$ *hyp*
3. $l = a_{\{1,2\}}$ 1, 2, transitivity of =
4. $l = r \rightarrow l = a$ 2 – 3, $\rightarrow I$

Thus, if we have a situation in which the most active dog in the world is Ramsey, then in that situation we also have the implication,

The laziest dog in the world is Ramsey implies that
the laziest dog is the most active dog.

This seems very odd. Recall our discussion regarding implication in chapter 1 above. There we said that in an implication we have to take the antecedent seriously. To use C. I. Lewis's phrase again, we have to give the paradoxer his or her hypothesis and see what follows from that. In this case, to evaluate the implication 'Ramsey's being the laziest dog in the world implies that the laziest dog is the most active dog' we have to set aside the fact that Ramsey is the world's most active dog and take seriously the hypothesis that he is the world's laziest dog.

One might object that our counterexample to the transitivity of identity relies on the use of definite descriptions rather than directly referring terms like proper names. This is true, but in this case it does not matter. For the phrase 'the cardinality of $\{1, 2\}$ ' that is used in the Hardy argument is also a definite description, and this is the argument that we wish to reject.

5.8 Two arguments concerning negation

In this section we look at two very similar arguments concerning the semantics of negation, the first due to Timothy Williamson and the second due to David Lewis. Williamson's argument purports to show that there is something incoherent in holding that there are truth value gaps and Lewis' argument tries to show that there is something incoherent about claiming that there are inconsistent situations.

To make present these arguments succinctly, I will use the notation ' $s \models A$ ' to mean ' A is true at s '. Both arguments use the following rule. For all formulae A and all situations s ,

$$(N) s \models \sim A \text{ iff } \text{Not } s \models A.$$

We shall discuss this rule later at some length.

Let us start with Williamson's argument. The argument is presented in (Williamson 1994), pp. 188–9. In our notation, we can rewrite Williamson's

argument as follows⁷:

- | | |
|--|---|
| 1. Not ($s \models A$ or $s \models \sim A$) | <i>hyp</i> (denial of bivalence for s) |
| 2. Not ($s \models A$ or Not $s \models A$) | 1, (N) |
| 3. Not $s \models A$ and Not Not $s \models A$ | 2, DeMorgan's law |
| 4. Not $s \models A$ and $s \models A$ | 3, Double negation elimination |

Line 4 is not necessary and Williamson does not include it in his argument, but it does make the contradiction more explicit.

Lewis' argument derives a similar conclusion from the premise that there are inconsistent situations. We quoted Lewis' argument in chapter 4 above, but we should repeat it here.

For comparison, suppose travellers told of a place in this world – a marvellous mountain, far away in the bush – where contradictions are true. Allegedly, we have truths of the form 'On the mountain both P and not P'. But if 'on the mountain' is a restricting modifier, which works by limiting domains of implicit and explicit quantification to a certain part of all that there is, then it has no effect on the truth-functional connectives. Then the order of modifier and connectives makes no difference. So 'On the mountain both P and Q' is equivalent to 'On the mountain P, and on the mountain Q'; likewise 'On the mountain not P' is equivalent to 'Not: on the mountain P'; putting these together, the alleged truth 'On the mountain both P and not P' is equivalent to the overt contradiction

⁷ Williamson's actual argument is couched in terms of utterances and the propositions that they express. My version captures the main moves in the framework of the present theory. I think that I have done no violence to Williamson's intentions.

But here is Williamson's original argument. Where u is an utterance and π is a proposition, Williamson proposes the following two biconditionals:

- (T) If u says that π , then u is true iff π
 (F) If u says that π , then u is false iff not π

Then he makes an assumption:

- (0) u says that π .

He now represents the denial of bivalence as

- (1) Not: either u is true or u is false.

From (T) and (0) by modus ponens we derive (2a) below:

- (2a) u is true iff π .

And from (F) and (0) we derive (2b), viz.,

- (2b) u is false iff not π .

So, from (0), (2a), and (2b) we obtain:

- (3) Not: either π or not π .

From (3) and DeMorgan's law, we derive

- (4) not π and not not π .

'On the mountain P, and not: on the mountain P'. That is there is no difference between a contradiction within the scope of a modifier and a plain contradiction that has the modifier within it. So to tell the alleged truth about the marvellously contradictory things that happen on the mountain is no different from contradicting yourself. But there is no subject matter, however marvellous, about which you can tell the truth by contradicting yourself. ((Lewis 1986), p. 7n)

Let us translate this talk of mountains into talk of situations. We then abstract the following argument:

- | | |
|--|---|
| 1. $s \models A \wedge \sim A$ | <i>hyp</i> (there is an inconsistent situation) |
| 2. $s \models A$ | 1, Truth Cond. for conjunction |
| 3. $s \models \sim A$ | 1, Truth Cond. for conjunction |
| 4. Not $s \models A$ | 3, (N) |
| 5. $s \models A$ and Not $s \models A$ | 2, 4 |

Once again, we have a contradiction.

We have adopted truth value gaps (in some situations) and the existence of inconsistent situations. Thus we need a response to these arguments.

Note that it does not help to claim that the problem is that the metalinguistic 'not' used in these principles is classical. For if we were to replace it with our relevant negation, the same arguments would hold. No principles used in these arguments violate relevant canons of proof.

Thus, I claim that the problem lies in the assumption that our object language negation ought to be interpreted in this straightforward way in terms of a metalinguistic negation. This assumption, however, is rather natural. It seems to work for some of our other connectives. We hold, for example, that,

$$s \models A \wedge B \text{ iff } s \models A \text{ and } s \models B.$$

Here we interpret our object language conjunction in terms of a metalinguistic conjunction. If this sort of correspondence is reasonable in the case of conjunction, why not in the case of negation?

The answer lies in the dual nature of situations. On the present theory, situations are the primary indices at which sentences are true or false. But they also provide the basis for a theory of *information*. I hold that a situation, s , makes a sentence 'A' true if and only if s contains the information that A. This is an informational form of Tarski's T-scheme. Now, it is not the case that a situation contains the information that A if and only if it does not contain the information that not-A. For consider the following example. Consider a situation that comprises all the information that I have about the actual world at the moment as I sit in my study in Wellington. Let us call this situation ' s '. I ask myself the following two questions. (1) Is it raining now in Toronto? (2) Does s contain

the information that it is raining in Toronto? Question (2) is easy to answer. s does not contain this information. But (1) has me stumped. I have no idea what the weather is in Toronto right now. This seems to me to be a very different question. Thus, it would seem that we should reject the right-to-left direct of (N). Let us call it 'Nrl', viz.,

(Nrl) If not $s \models A$, $s \models \sim A$.

Now, what about the other direction, i.e., Nlr, viz.,

(Nlr) If $s \models \sim A$, not $s \models A$?

To justify the rejection of Nlr it suffices to do two things. First, we should justify the acceptance in our model of inconsistent situations, i.e., situations that are incompatible with themselves. This justification takes place in various forms throughout this whole book. Second, we should show that there is an intelligible semantics of negation that does not entail Nlr. This alternative semantics, I have claimed, is that provided by the compatibility relation on situations.

So, we should reject the simple relationship between negation in our object language and negation in our semantic metalanguage. This might seem at first to be counterintuitive, but as soon as we understand the compatibility relation, our feeling of unease disappears.

5.9 Four-valued semantics

So far I have motivated the implicational approach to negation in its own terms. There is, however, an approach to negation that is growing in popularity among paraconsistent and relevant logicians and I should say a few words about what it is and why I am not adopting it.

This alternative approach is a four-valued semantics originally developed by Dunn (Dunn 1976) for the implication free fragment of the relevant logic R, and has been developed by Belnap who proposed it for use in artificial intelligence (Belnap 1977b), (Belnap 1977a). This semantics allows statements to take one of four values – True, False, Neither True nor False, Both True and False. Two of these values are designated – True and Both True and False. That is, both of these values are treated as 'true' values. They are treated as such because both include true. If a proposition is both true and false, then it is *at least true* (even if it is also false). We will represent a formula being at least true at a situation by ' $s \Vdash A$ '. The values False and Both True and False can be thought similarly as being *at least false* (even if one of them is also true). We represent a formula as being at least false at a situation by ' $s \dashv\vdash A$ '.

So, a formula is Both True and False at s if and only if

$$s \Vdash A \text{ and } s - \Vdash A.$$

It is True at s if and only if

$$s \Vdash A \text{ and } s \not\vdash A.$$

It is False at s if and only if

$$s \not\vdash A \text{ and } s - \Vdash A.$$

And it is Neither True nor False at s if and only if

$$s \not\vdash A \text{ and } s \not\vdash A.$$

In addition to the usual truth conditions for the various connectives, we have falsity conditions. The truth condition for conjunction is the usual,

$$s \Vdash A \wedge B \text{ iff } s \Vdash A \text{ and } s \Vdash B.$$

Its falsity condition is the following:

$$s - \Vdash A \wedge B \text{ iff } s - \Vdash A \text{ or } s - \Vdash B.$$

Similarly, the truth condition for disjunction is the standard one:

$$s \Vdash A \vee B \text{ iff } s \Vdash A \text{ or } s \Vdash B.$$

Its falsity condition is

$$s \not\vdash A \vee B \text{ iff } s - \Vdash A \text{ and } s - \Vdash B.$$

The treatment of negation in this semantics is quite elegant. Its truth condition is

$$s \Vdash \sim A \text{ iff } s - \Vdash A$$

and its falsity condition is

$$s - \Vdash \sim A \text{ iff } s \Vdash A.$$

In other words, $\sim A$ is true if and only if A is false and $\sim A$ is false if and only if A is true.

Compare this four-valued semantics to standard truth tables for classical logic. In a sense we get the same truth and falsity conditions in the classical case. The only difference here is that, in effect, a formula can take more than one truth value. Thus, it can be seen as a generalisation of classical truth tables.

This rather pretty semantics, as we have so far presented it, has a drawback from our point of view. It does not include truth conditions for implication. It was, however, extended to include implication by Richard Routley

((Routley 1984), also see (Routley *et al.* 1982). Routley did so by adopting the Routley–Meyer frame theory, removing the star operator, and adding a second ternary accessibility relation.

Routley’s four-valued frame contains a set of situations (some of which are designated, as in chapter 3 above), a ‘positive’ ternary relation R , and a ‘negative’ ternary relation, F . The truth condition for implication is the standard Routley–Meyer condition:

$$s \models A \rightarrow B \text{ iff } \forall x \forall y ((R s x y \ \& \ x \Vdash A) \supset y \Vdash B)$$

The falsity condition uses the negative accessibility relation:

$$s - \Vdash A \rightarrow B \text{ iff } \exists x \exists y (F s x y \ \& \ x - \Vdash B \ \& \ y - \nVdash A)$$

Thus, with the four-valued semantics, we have traded in the star operator for a second ternary relation.

The Routley continuation of Dunn’s semantics is both technically cumbersome and philosophically hard to understand. A more elegant semantics was developed in (Restall 1995). On this semantics, the falsity condition for implication is given in terms of the original Routley–Meyer ternary relation. An implication $A \rightarrow B$ is false at a situation s if and only if, there are situations t and u , such that $R t u s$, A is true at t , and $\sim B$ is true at u . Mares (forthcoming) further refines Restall’s idea and shows that a very simple semantics can be produced for the logic R . But both Restall’s semantics and my own require that we can make good philosophical sense of this falsity condition. Whether there is a good way of making sense of this falsity condition is a problem that I leave open.

5.10 Paraconsistency

A paraconsistent logic is a logic which does not make valid the rule of ex falso quodlibet, viz.,

$$\frac{A \quad \sim A}{\therefore B}$$

Relevant logic rejects ex falso in both forms:

$$\frac{A_\alpha \quad \sim A_\beta}{\therefore B_{\alpha \cup \beta}}$$

and

$$\frac{A_\alpha \quad \sim A_\alpha}{\therefore B_\alpha}.$$

Thus, relevant logic is a paraconsistent logic.

Philosophers and computer scientists have adopted paraconsistent logics for a variety of reasons. We can place ‘paraconsistentists’ on a spectrum, like a political spectrum.

On the right are those who think that paraconsistent logics are useful tools for dealing with inconsistent theories, inconsistent stories, and inconsistent sets of beliefs. We all have inconsistent beliefs. Our belief sets are too large and complicated for us to clean out old and inappropriate beliefs when we learn new things. Among these old beliefs there are often beliefs that contradict new beliefs or contradict one another.

In addition, in the history of science and mathematics we find useful and important but inconsistent theories. Niels Bohr’s early theory of the atom included features of the classical theory – his view predicted both that the electron would radiate all its energy and spin into the nucleus and that it wouldn’t. Isaac Newton’s use of his calculus in his proof of the product rule in the Introduction to his *Quadrature of Curves* is inconsistent. Throughout one proof, he assumes infinitesimal quantities are positive and divides through by them. At the end of the proof, however, he eliminates these quantities as if they had the value zero. To understand reasoning about inconsistent systems (or inconsistent reasoning about a system, as in Newton’s case), we need a paraconsistent logic.

On the left in our logical spectrum, are those who think that there are true contradictions. Most famous among the lefties are Graham Priest and Richard Sylvan. They believe that there are all sorts of contradictions. Consider the liar sentence ‘This sentence is false.’ According to Priest and Sylvan, the liar sentence is true but also false. Priest thinks that there are many true contradictions. There are, for example, sets that do and do not belong to themselves and things that are not identical with themselves.

We call those who think that there are true contradictions *dialetheists* and those who accept that there are interesting and useful inconsistent theories or inconsistent sets of belief, *doxastic paraconsistentists*. In this book, I am taking the position of a doxastic paraconsistentist.⁸ In the next section I will defend this position against an argument due to Priest.

⁸ Elsewhere I have taken an approach much closer to dialetheism (Mares 2000), but my current inclination is towards epistemic paraconsistency.

5.11 Priest's slippery slope

In (Priest 2000), Priest argues that every paraconsistentist should also be a dialetheist. His argument is in the form of a slippery slope. Priest distinguishes between three levels of paraconsistency, but we need only distinguish between two. For it is only the move from the second to the third level that threatens our own position.

Here is Priest's argument. Suppose that a person accepts that there are non-trivial, interesting, and useful inconsistent theories. That is, the person accepts the thesis of weak (or doxastic) paraconsistency, as we have defined it in the previous section. But, Priest asks, if there are interesting and useful theories that are inconsistent but non-trivial, then why shouldn't some of these theories be true? What Priest is asking for here is some principled objection to dialetheism, which principles are acceptable from the standpoint of doxastic paraconsistency.

Priest thinks that there is only one option open to the doxastic paraconsistentist – an appeal to the principle of consistency. Now, the doxastic paraconsistentist and the dialetheist disagree on whether the principle of consistency is correct. So, it would seem that the former needs an argument to support the acceptance of consistency.

Priest cites Aristotle's argument as the only 'sustained defence' of consistency in the philosophical literature. And he finds Aristotle's argument wanting. Thus, the doxastic paraconsistentist loses the battle and dialetheism triumphs.

Priest's argument uses a form of *onus pushing*. He claims that the burden of proof is on the doxastic paraconsistentist why she should not accept contradictions. The problem with onus pushing strategies in argumentation is that they are not very robust. Change perspectives slightly, and the onus can be pushed back to the other side. The present argument is a case in point. The fact that the principle of consistency is so firmly entrenched (both among philosophers and non-philosophers) gives us a good reason not to accept contradictions.

The doxastic paraconsistentist can claim that there have been useful inconsistent theories because each of these theories has approximated a consistent theory that has in turn approximated the truth. For example, Bohr's early theory of the atom was an inconsistent approximation of modern atomic theory. The original calculus approximated the modern theory of non-standard analysis, which in turn yields results that can be translated into theorems of a consistent version of the calculus.

This explanation of the success of these inconsistent theories, together with the deeply entrenched intuition that there are no true contradictions, in the absence of further argument, seems to give victory to the doxastic paraconsistentist.

5.12 Boolean Negation?

Our remarks in section 5.7 about intuitionist negation raise a further question. Should we add Boolean negation to our logic? The answer, I think, is 'no'. To see why I say this, let us first see what Boolean negation is.

Perhaps the easiest way to understand Boolean negation is through its truth condition. Let ' $\bar{A}$ ' be the Boolean negation of ' A '.

' $\bar{A}$ ' is true at s if and only if ' A ' fails to be true at s .

This truth condition seems at first glance both to give a clear meaning to the connective and to make it legitimate. Consider for a moment the truth condition for conjunction:

' $A \wedge B$ ' is true at s if and only if ' A ' is true at s and ' B ' is true at s .

Here conjunction is understood in terms of the conjunction of the semantic metalanguage. To put the point in terms of a slogan '*and*' means '*and*'. '*And*' should mean '*and*'. If it meant anything else, we would be appalled. The fact that we can interpret a connective in terms of the corresponding connective in the metalanguage provides a check on the theory of meaning that we are using. We want these correspondences. It would seem then that Boolean negation has passed the test. Boolean negation is understood in terms of metalinguistic negation.

There is, however, a problem here. It has to do with the partiality of information in our situations. We maintain that all information is persistent, that is, for any proposition A , if ' A ' is true at s and $s \sqsubseteq t$ then ' A ' is true at t . But, if $s \sqsubseteq t$ and ' B ' is true at t but ' B ' fails to be true at s , then ' $\bar{B}$ ' is true at s but not true at t . So the Boolean negation of a proposition is not always a persistent piece of information. In other words, in our system, the Boolean negation of a proposition is not always a proposition. Thus, we cannot include it as a connective in our scheme.

There are ways of modifying our semantics to include Boolean negation. One, due to Meyer and Routley, is to collapse our hereditariness relation into identity. That is, we set $s \sqsubseteq t$ if and only if $s = t$ (Meyer and Routley 1973). This is an adequate device from a technical point of view, but I cannot include it in my semantics without radically changing my view of situations and their role in reasoning. If I were to adopt Boolean negation and the Meyer–Routley semantics for it, I could not hold both that (1) a sentence is true in a world if and only if it is true in some situation that obtains at that world and (2) for some sentence ' A ' there is a possible world w and situations s and t such that both s and t obtain at w and ' A ' is true in s but not in t . For then I would have to hold that both ' A ' and ' $\bar{A}$ ' are true in w . To adopt this view, then, would be

to adopt a radically non-classical view of Boolean negation and reject the real motivation behind it.⁹

5.13 Denial

One feature of Boolean negation is that it is a form of *denial* negation. If ‘ $\bar{A}$ ’ is true, it precludes ‘ A ’ from being true. Sometimes we want to say something that precludes a proposition’s being true. Our relevant negation might appear to be too weak for this purpose. For there are situations in which both a proposition and its relevant negation are true. Thus, relevant negation is not a denial negation.

One route that non-classical logicians take to include denial in their theory is to add it as a type of speech act. I follow these logicians and add denial as a speech act and *rejection* as the propositional attitude that is expressed by denial.

Denial is a type of illocutionary force. It should not be confused either with the assertion of a negation or illocutionary negation. As Austin, Searle, and others have argued, a speech act has two components – a propositional content and an illocutionary force. Thus, a speech act has the form:

$$F(p)$$

where F is its force and p its propositional content. Our object language contains only formulae that express propositional contents.

Asserting is one sort of illocutionary force. The following sentence is an assertion of a negation:

$$\text{Ramsey is not barking at the neighbours.} \quad (5.1)$$

It can be easily distinguished from a sentence that contains an illocutionary negation such as

$$\text{I do not assert that Ramsey is barking at the neighbours.} \quad (5.2)$$

Sentence (5.1) is about what Ramsey is or isn’t doing. Sentence (5.2), on the other hand, is about what the speaker is or isn’t doing. In terms of the logical form of speech acts, (5.1) has the form $F(\sim p)$ and the latter has the form $\sim F(p)$.¹⁰

⁹ An anonymous referee argued that the real reason why we do not include Boolean negation in the language of relevant logic is that, if we do, we make valid $(p \wedge \bar{p}) \rightarrow q$, which is a paradox of strict implication. This is true. The original motive for barring Boolean negation was to get rid of this and other similar paradoxes. But my point here is that we can reach the same conclusion from the point of view of the present interpretation of the semantics for relevant logic.

¹⁰ The notion of illocutionary negation is due to John Searle ((Searle 1969), p. 32) and R. M. Hare (Hare 1970). Hare calls it ‘external negation’.

Denial, however, is quite different from either the assertion of a negation or illocutionary negation. Consider the following example due to Terence Parsons:

Paul Bunyan is not bald. (5.3)

A speaker might mean one of two things in saying (5.3). First, she might be claiming that Paul Bunyan has hair. Second, she might be saying that 'Paul Bunyan is bald' is not true. This latter claim does not amount to the claim that the statement is false, for she might also hold that Predications about nonentities have no truth value. As Parsons points out:

I might say 'Paul Bunyan is *not* bald' without thereby committing myself to the truth of the sentence 'Paul Bunyan is not bald', for I might think (as many people do think) that this sentence lacks truth value. I may only want to reject the sentence 'Paul Bunyan is bald' (or perhaps the proposition that it expresses), for I may think that both claims – 'Paul Bunyan is bald' and 'Paul Bunyan is not bald' – lack truth value. If I am told that I can only deny the former by asserting the latter, then that is a language game that I will not want to play. And I do not think that this is how ordinary English works. Sometimes saying a sentence with a negative word in a certain tone of voice just counts as a rejection of the corresponding positive version. ((Parsons 1984), p. 139)

I think that Parsons is right. Even if we think that 'Paul Bunyan is bald' has a truth value, we understand what someone means when she rejects the claim that it does. This is a denial.

Philosophers who accept truth-value gaps, such as many-valued logicians and supervaluationists, must accept that denial is not reducible to the assertion of a negation. Moreover I hold with Parsons that everyone accept denial regardless of their other beliefs about logic. If we hear someone saying, for example 'Paul Bunyan is not bald and he is not not bald either', we usually do not attribute to her the belief that there is a true contradiction. Rather, we take her to be denying both 'Paul Bunyan is bald' and 'Paul Bunyan is not bald'. We understand what this speaker is saying whether we agree with her position on the logic of non-referring terms or not.

The circumstances become more complex if we add to our language a propositional operator that means '*is accurately denied*'. It would seem that we then have Boolean negation in our language. For, in any situation *s*, *A* is accurately denied if and only if *A* fails to be true. The problem here is the same as the problem with including Boolean negation in our language. The 'information' that a proposition is accurately denied is not persistent, and so cannot count as information at all in our sense. But we can include a similar binary operator, ' δ ', that means 'is accurately denied in'. We can say in our language that a proposition is (or is not) accurately denied in a particular situation. Let us let ' $\delta(A, s)$ ' mean that '*A*' is accurately denied in *s*. Now, this statement is not

at all the same as the Boolean negation of 'A'. For it is far from clear that we should include the truth condition,

' $\delta(A, s)$ ' is true at s iff 'A' fails to be true at s .

Although there is a sense in which a situation 'knows' what it makes true, it is less obvious that it knows what it fails to make true. What it makes true is determined by the information *in* that situation. But what it fails to make true is determined by information *about* that situation that may not also be in that situation. Given the sort of semantics we are creating, we do not want situations to be 'omniscient' about themselves. Thus, we reject the above truth condition and so cannot recover Boolean negation by way of denial.

If we do add the operator ' δ ' to our language, however, we may have to change our semantics in an important way. We cannot have any possible situations that are bivalent. For suppose that s is a worldly situation. Also suppose that we have the resources in our semantics that allow a sentence to refer to itself. Let ' Δ ' mean 'this sentence is accurately denied in s '. Then, ' Δ ' is equivalent to ' $\delta(\Delta, s)$ '. By the truth condition given above, we have

' $\delta(\Delta, s)$ ' is true in s iff ' Δ ' fails to be true in s .

But, ' Δ ' is equivalent to ' $\delta(\Delta, s)$ ', so we also have

' Δ ' fails to be true in s iff ' $\delta(\Delta, s)$ ' fails to be true in s .

Thus, we can derive

' $\delta(\Delta, s)$ ' is true in s iff ' $\delta(\Delta, s)$ ' fails to be true in s .

So, if we have both self-reference and a denial operator we have to remove all worldly situations from our semantics. This means that we must get rid of worldly situations and with their expulsion we would have to reject the law of excluded middle. We have already assumed that a logical situation obtains at each possible world. If a logical situation obtains at a possible world and makes true the law of excluded middle, then that situation is a worldly situation. For any situation that extends it would be inconsistent.

We could, alternatively, banish denial operators from our language and accept a Tarski-like theory in which they can be contained only in metalanguages. This would allow us to retain worldly situations and the tautologies of classical logic.

With these few brief and inconclusive remarks we will leave the topics of how or whether we should allow denial to be expressed in our language. We will assume for the rest of this book that we do not have denial operators. This assumption makes things much easier.

6 Modality, entailment and quantification

6.1 Relevance and modality

In chapter 1 above, we saw that relevant logicians distinguish between implication and entailment. Implication, as we have seen, is a contingent relation between propositions. We have explicated this notion using the theory of situated inference. Now we turn to the topic of entailment. We introduce this topic historically.

For the beginning of this history, we return to the work of C. I. Lewis. Recall from the introductory chapter above that he was dissatisfied with the treatment of implication as material implication. In order to avoid the paradoxes of material implication, he introduced his own theory of implication, that was supposed to be codified in the *strict implication* of his various modal logics (among which S2 was supposed to give the most accurate codification). The addition of necessity to the definition of implication was supposed to make implication a relation between the meanings of statements rather than their truth values. In so doing, Lewis produced a theory of what we call ‘entailment’ rather than a theory of implication, and we treat it as a theory of entailment in this chapter.

In the paper that began the study of relevant entailment, ‘Begründung einer strengen Implikation’, Wilhelm Ackermann criticises Lewis’ systems. In particular, Ackermann rejects *ex falso quodlibet*, i.e.,

$$(A \wedge \sim A) \rightarrow B,$$

as a general truth about entailment. To say that *A* entails *B*, according to Ackermann, is to hold that ‘a logical connection holds between *A* and *B*’. But there is no real logical connection between $A \wedge \sim A$ and *B* (Ackermann 1956, p. 113).¹

In providing a system that does not make Lewis’ error, Ackermann retained the idea that entailment is a logical relationship between the meanings (*Inhalt*) of statements. He viewed this relationship as one of containment. He says that

¹ Like Lewis, Ackermann uses ‘implication’ to refer to what we call ‘entailment’. To avoid confusion, I use ‘entailment’ where Ackermann uses ‘implication’.

' $A \rightarrow B$ ' says that the 'content (*Inhalt*) of B is part of the content of A ' (ibid.). No theory of the content of statements is given by Ackermann, so it is difficult to assess either this claim or the success of his system at codifying it.

Ackermann's system of entailment – called Π' – was reformulated by Alan Anderson and Nuel Belnap as their logic E, for 'entailment'. But E has had an unfortunate history.

6.2 The sad story of E

The entailment connective of E, as Anderson and Belnap viewed it, was supposed to be a relevant strict implication. In the early 1960s Anderson and Belnap also formulated the logic R of relevant implication, the logic that we have been studying. Thus, if R is supposed to capture relevant implication, and E is supposed to capture strict relevant implication, then it would stand to reason that we can capture the notion of entailment as formulated by E in R expanded to include modal operators.

Thus Bob Meyer added a necessity operator to R to create the logic NR. And he conjectured that this new system captured what Ackermann wanted his system Π' to formalise and Anderson and Belnap wanted from their logic E. At the time, Meyer thought that the conjectured equivalence of these systems would show that all three of them had got the notion of entailment right. In 1971, Meyer said that:

Within the limits of experimental error it can now be reported that the Anderson–Belnap system E of entailment, identical in its stock of theorems and in other significant respects (as it has turned out) with the earlier systems of *strengte Implikation* of W. Ackermann, furnishes a true and correct formal counterpart to the intuitive notion of entailment. One more philosophical problem, you will be happy to know, has been definitively and finally solved; anyone who might have been tempted to work on it is referred instead to the mind–body problem, which if we all pull together ought to be disposed of shortly. ((Meyer 1971), p. 810)

The logics NR and E do in a sense capture the same notion of entailment. Suppose that we define in NR strict implication as follows:

$$A \multimap B =_{df} \Box(A \rightarrow B).$$

Then the class of theorems that we can formulate using only ' $\multimap$ ', propositional variables, and parentheses in NR is the same as the class of theorems that we can prove in E in the vocabulary of entailment, propositional variables, and parentheses.

But there are important differences between the two logics. Larisa Maksimova showed that, in the vocabulary of conjunction, disjunction, negation, and strict implication, NR contains a theorem that is not a theorem of

E (Maksimova 1971). Thus, there is a sense in which the entailment of E and the strict implication of NR are not the same.

There are various ways we can respond to this divergence of E and NR. First, we can continue to believe that E is 'the system' of strict relevant implication and reject NR. Second, one might abandon E in this role and look at logics like NR that explicitly combine R with some form of modality. Third, we might find another use for either R or E that renders the connection (or lack of connection) between the two irrelevant.

I have already provided an independent interpretation and motivation for R. So, the question for us at this stage is the following. *In the context of the theory of situated inference, what should we make of the notion of entailment?* We begin to answer this question by looking at the question, 'What is entailment?'

6.3 What is entailment?

The classic source of the philosophical notion of entailment is G. E. Moore's essay 'External and Internal Relations'. Moore introduces the term 'entailment' in the following passage from that essay:

We require ... some term to express the *converse* of that relation which we assert to hold between a particular proposition *q* and a particular proposition *p*, when we assert that *q* follows from or is deducible from *p*. Let us use the term 'entails' to express the converse of this relation. We shall then be able to say truly that '*p* entails *q*,' when and only when we are able to say that '*q* follows from *p*' or 'is deducible from *p*', in the sense that a conclusion of a syllogism in Barbara follows from the two premises, taken as one conjunctive proposition; or in which the proposition 'This is coloured' follows from 'This is red.' '*p* entails *q*' will be related to '*q* follows from *p*' in the same way in which 'A is greater than B' is related to 'B is less than A'. ((Moore 1922), p. 291)

The main ideas here are that the entailment is a relation between propositions and that it is the converse of the relation of deducibility.

This concept of entailment has remained in philosophical tradition. For example, Hughes and Cresswell say

An important modal notion is that of entailment. By this we understand the converse of the relation of following logically from (when this is understood as a relation between propositions, not wff) i.e. to say that a proposition, *p*, entails a proposition, *q*, is simply an alternative way of saying that *q* follows logically from *p*, or that the inference from *p* to *q* is logically valid. ((Hughes and Cresswell 1996), p. 203; see (Hughes and Cresswell 1968), p. 23)

Both Moore and Hughes and Cresswell, then, take entailment to be the converse of deducibility.

Later in ‘External and Internal Relations’, Moore gives an apparently semantic characterisation of entailment in terms of possible worlds (p. 293). In discussing the proposition $x = A$ entails xP , Moore says

This last proposition again, is, so far as I can see, either identical with or logically equivalent to the propositions expressed by ‘anything which were identical to A would, in any conceivable universe, necessarily have P ’ or by ‘ A could not have existed in any possible world without having P ’; just as the proposition expressed by ‘In any possible world a right angle must be an angle: is, I take it, either identical with or logically equivalent to the proposition ‘(x is a right angle) entails (x is an angle)’. ((Moore 1922) p. 293)

Although it would be anachronistic to attribute a real understanding of the distinction between semantics and proof theory to Moore in 1919 (the year his article was published), the above characterisation looks like a modern semantic characterisation of entailment.

Both Moore and Hughes and Cresswell treat entailment as a form of strict implication and we will follow their lead in this. These philosophers, in effect, treat entailment in a strict implication in the sense of S5. Hughes and Cresswell take entailment to be logically necessary material implication. A logical necessity, for Hughes and Cresswell, is something that is true in all possible worlds ((Hughes and Cresswell 1968), p. 22 and (Hughes and Cresswell 1996), p. 203).²

We shall look at S4-like and S5-like modalities soon. But before I get to that and to alternative views of entailment, I would like to deal with an issue that has plagued discussions of entailment and implication for some years.

6.4 Are we confusing use and mention?

Consider Moore’s definition of ‘entailment’. We say that A entails B if and only if B is deducible from A . One natural way of reading this definition is to take entailment to be a relation between sentences. Thus, ‘ A entails B ’ if and only if we can derive the sentence ‘ B ’ from the sentence ‘ A ’ in some appropriate formal system. We will write this as “ $A \vdash B$ ”. This statement, it would seem, is in the metalanguage of the formal system.

Now consider a statement about a nested entailment. Let’s express the claim that A ’s entailing B entails C ’s entailing D in a perspicuous way. It would seem that we get

$$“A \vdash B” \vdash “C \vdash D”.$$

² Cresswell has told me in conversation that although we are safe in attributing this view to him, Hughes was not as clear about his view on the matter.

This statement is in the meta-metalanguage of our formal system. And the middle turnstile does not have the same meaning as the turnstiles in the expressions flanking it. Thus, W. V. O. Quine for one, thinks talk of nested entailments and the treatment of entailment as a connective in a formal language is a mistake ((Quine 1950) pp. 50f.). Once we straighten out where the quotation marks should go, we find that our alleged nested implications are in fact claims in a metalanguage (or meta-metalanguage, meta-meta-metalanguage, etc.). Thus, the idea of formalising a logic of entailment is misconceived.

In reply to Quine, Alasdair Urquhart has argued that even if we think that entailment is a relation between sentences, we need not think that we must climb the hierarchy of metalanguages in this way. Suppose that our language contains names for the sentences of the language itself. In this language we can have a binary predicate '*E*' such '*E*(*x*, *y*)' means that the sentence named by *x* entails the sentence named by *y*. Let $\ulcorner A \urcorner$ be a name of the sentence '*A*'. Then we can express '*A*'s entailing *B* entails *C*'s entailing *D*' in our language as ' $E(\ulcorner E(\ulcorner E(\ulcorner A \urcorner, \ulcorner B \urcorner) \urcorner, \ulcorner C \urcorner) \urcorner, \ulcorner D \urcorner) \urcorner$ ' (Urquhart 1981).

As Urquhart notes, we need not stop here. Consider logics that contain a provability operator, such as the logic GL (see (Boolos 1993)). These logics represent Gödel's provability *predicate* as a modal operator. Provability predicates are monadic predicates of sentences (or rather, sentence names) (see chapter 11 below). The representation of these predicates in modal logics seems accurate, so why not do the same thing here and represent a binary predicate on sentence names as a binary operator in a modal-like logic? This is exactly what we try to do when we are creating a logic of entailment.

6.5 The metaphysics of modality

We already have an interpretation of relevant implication. Thus, to understand strict relevant implication we need a theory of necessity.

Semantically we treat modality in the standard Kripkean sense. A situation can 'see' or 'has accessible to it' a number of situations. ' $\Box A$ ' holds in *s* if and only if in each of the situations accessible to *s* (for some given binary accessibility relation) '*A*' is true, . If and only if '*A*' is true in at least one accessible situation, ' $\Diamond A$ ' is true in *s*. Let us call our modal accessibility relation *M* (to distinguish it from the ternary implicational accessibility relation *R*). Thus, our truth conditions for the modal operators written formally look like this:

' $\Box A$ ' is true in *s* if and only if $\forall x(Msx \supset 'A' \text{ is true in } x)$

' $\Diamond A$ ' is true in *s* if and only if $\exists x(Msx \ \& \ 'A' \text{ is true in } x)$

Moreover, on our treatment, the usual relationship holds between $\Box$ and $\Diamond$, is

valid in our semantics, viz.,

$$\sim\Box\sim A \leftrightarrow \Diamond A.$$

Thus, formally speaking, we have a reasonably standard treatment of modality.

Philosophically, we have a problem that is shared among all metaphysical treatments of modality. We need to answer the question, what determines which situations are modally accessible to which other situations? In what follows, I will answer this question for at least the few sorts of modality that we will use in the other chapters of this book.

The treatment of necessity and possibility that I adopt is a slight modification of Agelika Kratzer's semantics for 'can' and 'must' (Kratzer 1977). On Kratzer's semantics, the sentence

The ancestors of the Maori *must* have arrived from Tahiti.

should be understood as expressing roughly the same proposition as

In view of what is known, the ancestors of the Maori *must*
have arrived from Tahiti. (ibid., pp. 338–40)

The word 'must' here does not express an 'absolute' necessity. Rather, 'must' has the role of telling us what is implied by our knowledge of Maori history.

Kratzer represents the role of 'must' in sentences like the above by treating it as a relative modal phrase. She represents it more perspicuously as:

Must_in_view_of_what_is_known,
the_ancestors_of_the_Maori_came_from_Tahiti) (see ibid., p. 341)

In possible world semantics, the treatment of relative modal phrases is reasonably intuitive. The first argument of the relative modal operator, that is, 'what is known', determines a set of propositions, those propositions about the Maori that are currently known. This set of propositions, in turn, determines a set of possible worlds, that is, those worlds that satisfy all those propositions. The sentence given above is true if and only if the second argument of the relative modal phrase is true at all the worlds determined by the first argument.

Most semanticists, I think, would agree with Kratzer that in standard usage, 'must' is a restricted modal operator, that is, it does not quantify over all possible worlds. On the approach to modality that I will develop here, all modal operators are restricted, in the sense that they do not quantify over all situations. For example, metaphysical necessity is a modality that philosophers usually treat as unrestricted; they hold that a proposition is metaphysically necessary if and only if it is true in all possible worlds. I follow them, to some extent in this treatment of metaphysical necessity. As we shall see, a proposition is metaphysically necessary *in a worldly situation* if and only if

it is true in all worldly situations. But not all situations are worldly (hence metaphysical necessity is not unrestricted) and this biconditional does not hold in general when the context of evaluation is not a worldly situation. Thus, even metaphysical necessity on the present view is not an unrestricted modality.

The mechanism which determines modal accessibility relations in our theory is very much like the one Kratzer employs. In this chapter I will discuss only alethic modalities, and only three of them – nomic necessity, metaphysical necessity, and a sort of necessity connected with the notion of entailment.

First we will look at the basic formal properties of accessibility relations. Then we will use the example of nomic necessity to demonstrate how our metaphysics of modality operates.

We use binary accessibility relations to give truth conditions for our modal operators, just as in possible world semantics for modal logic. Let M be a modal accessibility relation. Then there will be some necessity operator, $\Box$, such that the following truth condition holds:

‘ $\Box A$ ’ is true in s if and only if $\forall x(Msx \supset \text{‘}A\text{’ is true in } x)$.

The relation M will also satisfy certain formal postulates, all of which are listed in appendix B below. Here we will look at only one such postulate, viz.,

If Mst , then Ms^*t^* .

This postulate makes true the following truth condition:

‘ $\Diamond A$ ’ is true in s if and only if $\exists x(Msx \ \& \ \text{‘}A\text{’ is true in } x)$,

where ‘ $\Diamond A$ ’ is defined as ‘ $\sim \Box \sim A$ ’.

Certain accessibility relations will satisfy special postulates. For example, if M is supposed to characterise a *normal* modality, then it will satisfy the following postulate:

If s is logical and Mst , then t is also logical.

This postulate makes valid the rule of necessitation, that is,

$$\frac{\vdash A}{\therefore \vdash \Box A}.$$

Any regular modal logic that contains the rule of necessitation is called a *normal modal logic*.³

³ Here ‘regular modal logic’ is to be interpreted in *relevant* terms, as allowing the following rule to be derived:

$$\frac{\vdash A \rightarrow B}{\vdash \Box A \rightarrow \Box B}$$

Now let us consider nomic necessity. The treatment of nomic necessity is rather straightforward and it uses only devices that we have already introduced in previous chapters, and so it makes a good starting point for our discussion of modality.

Instead of employing propositions as Kratzer does, we define a nomic accessibility relation using SOA. As we said in chapter 3 above, there is a class of SOA that represent laws of nature. Let us call these 'lawlike SOA'. Now consider an arbitrary situation, s . In order to determine which situations are nomically accessible to s , we first take the set of situations which contain all of the lawlike SOA that are contained in s . Let's call this set $N(s)$. We then set Mst if and only if t is in $N(s)$ and t^* is in $N(s^*)$. Clearly, M satisfies the postulate given above.⁴

We can also see that this nomic accessibility relation is reflexive and transitive. As is the case for modal logics based on the classical propositional calculus, the reflexivity of the accessibility relation implies that our nomic necessity relation obeys the T axiom, viz.,

$$\Box A \rightarrow A.$$

In addition, since nomic accessibility is transitive, nomic necessity also obeys the 4 axiom:

$$\Box A \rightarrow \Box \Box A.$$

So far, the logic of nomic necessity seems pretty standard. It seems more or less like a relevant version of S4. But there is a big difference between this system and S4. The logic of nomic necessity is not a normal modal logic. That is, its class of theorems is not closed under the rule of necessitation.

Consider a theorem of R, say, the law of excluded middle, $A \vee \sim A$. It seems odd to say that it is nomically necessary that excluded middle holds. Taking the logic of nomic necessity to be a non-normal system allows us strongly to distinguish between nomic necessities and metaphysical necessities; the latter need not be a subset of the former.

We can generalise this approach to treat a variety of different modalities. Like nomic necessity, some modalities are strongly associated with a class of SOA. When this is the case, we can define two modal accessibility relations. First, we can start, as we did above, with a situation s and the set of situations that contain all the salient SOA that are in s . Second, we can start with a situation s and the set of situations s' that contain all the salient SOA that are in s and if s' contains a SOA σ of the salient type then σ is also in s . This second approach will give us an accessibility relation that is symmetrical.

⁴ We also conjecture that M satisfies the other 'housekeeping' postulate, if $s \leq s'$ and $Ms't'$, then there is some $t \leq t'$ such that Mst .

Which of these relations we choose to model a given modal operator depends on how we use that operator. We consult speakers' intuitions on how the operator is used to determine whether or not we want a symmetrical accessibility relation. To make this determination we can ask, for example, whether the operator obeys the Browerische axiom, viz.,

$$p \rightarrow \Box \Diamond p.$$

If speakers have the firm intuition that the Browerische axiom holds, then we accept the symmetrical accessibility relation. If they have the intuition that it does not, then we accept the non-symmetrical relation. Of course, in many cases, speakers will not have firm intuitions. In such cases, we can claim what they mean using the operator is underdetermined, or even ambiguous.

6.6 The converse of deducibility

We now return to the problem of entailment. We will say that A entails B in a situation s if and only if, in s , it is logically necessary that $A \rightarrow B$. In order to understand what this truth condition means here we need to know what 'logically necessary' means here.

Again we couch our theory of a sort of necessity in terms of what situations are accessible to which other situations. Like other sorts of modal accessibility, logical necessity is determined by a sort of information. A situation s can see a situation t if and only if the information that can be used in a purely logical inference in s is also contained in t . At a logical situation, for example, perhaps only the information from the class of models of implication can be used. In other situations, the class of models may not be present, have only some of the properties that it has in logical situations, or have some properties that it doesn't have in logical situations. As in the case of nomic necessity, we can derive two sorts of accessibility from this theory. We can have $M_L st$ if and only if t contains all the 'logical' information that is present in s . And $M'_L st$ if and only if t contains exactly the same logical information that is present in s .

Now, all we have to do is choose one of these accessibility relations and use it to give a truth condition for entailment. Using M_L , we first give the standard truth condition for the necessity associated with it, that is,

$$\text{'}\Box_L A\text{' is true in } s \text{ if and only if } \forall x (M_L sx \supset \text{'}A\text{' is true in } x).$$

Given this truth condition and the following definition of entailment, we can derive the truth condition for that connective:

$$A \rhd B =_{df} \Box_L (A \rightarrow B).$$

The truth condition for entailment is

$$s \models A \rightarrow B \text{ if and only if } \forall x(M_L s x \supset \forall y \forall z((Rxyz \ \& \ 'A' \text{ is true in } y) \supset 'B' \text{ is true in } z)).$$

We can simplify this truth condition by defining a ternary accessibility relation for entailment. This relation is defined as follows:

$$Estu \text{ if and only if } \exists x(M_L s x \ \& \ Rxtu).$$

Then we can write the truth condition for entailment as follows:

$$s \models A \rightarrow B \text{ if and only if } \forall x \forall y((Estxy \ \& \ 'A' \text{ is true in } x) \supset 'B' \text{ is true in } y).$$

Thus we have a truth condition for entailment that mirrors the condition for implication.

Using Kratzer's method we can generalise this notion of entailment. There are many different notions of deducibility. There is deducibility in standard mathematics, there is a constructive notion of deducibility, and so on. By a notion of deducibility here, I mean a pre-formal notion. I do not mean 'deducibility in intuitionist logic', 'deducibility in classical logic', and so on. Rather, I want to capture more intuitive, pre-logical notions. Using Kratzer's method, we can isolate those SOA in a situation s that determine a notion of deducibility. These SOA will determine a set of situations accessible from s . Thus, we can formalise a range of notions of entailment in this system.

6.7 Metaphysical necessity

Another sort of modality that I would like to treat is the philosopher's favourite sort of modality – metaphysical necessity. And we will need it for the theory of conditional presented in chapter 7 below.

In what follows, I sketch a theory of metaphysical necessity based on a view of Bill Lycan and Stuart Shapiro (Lycan and Shapiro 1986). In that paper, Lycan and Shapiro develop an ersatzist theory of possible worlds. In order to construct a set of possible worlds, they require that the set of essential truths be stipulated before the construction can begin (ibid., p. 347). According to Lycan and Shapiro, all truths concerning only abstract objects are among the metaphysically essential truths, as are all 'purported metaphysical laws' (e.g., 'Everything is either material or ideal', 'Time travel does not occur'), and perhaps some features of individuals. These features of individuals might include what species they are (e.g. 'Ramsey is a dog') or their origin.

This might seem like a rather vague list, but as Lycan and Shapiro point out every account of possibility requires ‘stipulation of this sort’. They say:

So far as we can determine, the only way to specify a certain relative modality by reference to a special type of world is to stipulate the set of propositions distinctively contained by every world of that type. If so, stipulation of this sort is essential to *any* account of possible worlds. (Ibid.)

One might think that Lycan and Shapiro lose the advantage of a non-ersatzist account of possible worlds (like David Lewis’) since these do not need to specify a set of metaphysical truths in order to determine their non-relative form of necessity, that is, to determine which worlds are metaphysically possible, since these are taken as primitive to the theory. But Lycan and Shapiro disagree. Suppose, for example, two people are disputing whether there is a world at which ‘Socrates is crustacean rather than human.’ To adjudicate this question, one must rely on his or her knowledge of what is metaphysically necessary. This cannot come from the mere existence of a set of possible worlds. Thus, if we do not want to fall into complete scepticism about metaphysical modality, we have to admit that we are able (more or less) to stipulate what sorts of truths count as metaphysically necessary.

Note that Lycan and Shapiro are not claiming that we know all the metaphysically necessary truths. But they are claiming that we can know what sorts of truths count as metaphysically necessary. I agree. I think that we can stipulate what sorts of SOA determine metaphysical necessity.

When we look at what situations are metaphysically accessible to a given situation, we gather together the SOA of the sorts that determine metaphysical possibility. Then we see which situations also have those SOA. We apply our formal criteria to them and so determine which are metaphysically accessible.

An adequacy criterion for a theory of metaphysical possibility is that it make truth the following claim. *If s and t are both worldly situations, then s and t are metaphysically accessible from one another.* Clearly there are many accessibility relations that satisfy this criterion, so it can be taken only as a necessary and not as a sufficient condition for the correctness of the choice of a relation.

6.8 Fallacies of modality

Before the days of Kripke and Putnam, it was a commonplace in philosophy that a statement about a contingency could not entail a proposition about what is necessary. Relevance logicians call entailments of this sort, ‘fallacies of modality’.

We won’t be very rigorous about the definition of a fallacy of modality, since the exact definition is pretty tricky. But we can get the general idea fairly easily.

First, we need the notion of a necessitive. A necessitive is a proposition that says that something is necessary. ' $\Box p$ ' is a necessitive, and so is ' $q \wedge \Box p$ ', but ' $p \vee \Box q$ ' is not, since the latter can be true without q 's being necessary. *A statement commits a fallacy of modality if and only if it says that a necessitive is entailed by some non-necessitive.*

The issue of fallacies of modality was first raised in the relevant logic literature in Ackermann's paper discussed above (Ackermann 1956). Ackermann, however, did not claim that having a 'modally fallacious' theorem in one's system was bad in itself. He proved that his system does not contain any modally fallacious theorems in order to prove something else. Namely, he used this proof to show two things. First, he showed that certain relevantly fallacious schemes are not theorems of his system, such as ' $A \rightarrow (B \rightarrow B)$ ', ' $(A \wedge \sim A) \rightarrow B$ ', and ' $A \rightarrow (A \rightarrow A)$ '. Second, he showed that his system does not contain certain theorem schemes that C. I. Lewis also rejected, for example, ' $A \rightarrow ((A \rightarrow B) \rightarrow B)$ ' (ibid., p. 127).

From a philosophical point of view, Alberto Coffa has given us an argument that we should reject all modally fallacious statements. For example, Coffa motivates this rejection by appealing to an epistemological principle that he calls 'The Platonic Principle'. The Platonic Principle states that 'necessary knowledge cannot derive from experience' ((Anderson and Belnap 1975) p. 245, §22.1.2). As Coffa notes, the Platonic Principle has been held by philosophers from a variety of philosophical traditions (ibid.). Plato's argument for the doctrine of recollection in his dialogue *Meno* appeals to the Platonic Principle, as does Saint Augustine's argument for his theory of illumination, and the rationalists' arguments for innate ideas. And Hume's view that we cannot derive knowledge of relations of ideas from matters of fact alone (*Treatise* Bk III, Pt I, Section I) is also a variant of the Platonic Principle.

But there are many contemporary philosophers who reject Plato's principle. Hilary Putnam, for example, has argued that whereas we have discovered that water is H_2O empirically, it is a necessary truth. I do not want to try to decide this issue here. But, whatever we think about Plato's principle, we can see that it is a virtue of relevant modal logic that it does not force any such 'fallacious' statements upon us. Relevant modal logics give us this ability, whereas classically based normal modal logics do not. Every normal modal logic based on classical logic contains the rule,

$$\frac{\vdash A}{\therefore \vdash B \rightarrow \Box A},$$

where ' $\rightarrow$ ' is material implication. Thus, every classical normal modal logic commits 'fallacies of modality'. This rule, of course, is not valid in the relevant modal logics.

Thus, adopting relevant modal logic gives us a type of freedom. We can formulate logics that do not commit fallacies of modality. But we can also say that there are sorts of modality that do allow the deduction of necessitives from non-necessitives.

6.9 Deontic logic and Hume's principle

Deontic logics are systems that include the operator 'It ought to be that', usually formalised by ' O '. This operator is usually treated as a modal operator. In the semantics for standard deontic logics, a world w is related to a set of worlds that are morally ideal from the standpoint of w . The formula OA is said to be true at a world w if and only if A is true in every world ideal from the standpoint of w .

If we add the 'ought' operator to relevant logic, we can adopt the same semantical insight. That is, for each situation s we set the situations that are morally ideal from the point of view of s accessible to s . As we saw with regard to nomic modality in section 6.5 above, we need not make this logic a normal modal logic. Once again, consider the law of excluded middle, $A \vee \sim A$. It would seem counterintuitive that it *morally ought to be the case* that $A \vee \sim A$ for every proposition A . Thus, it would seem that we should take relevant deontic logic to be a non-normal modal logic.

We have claimed that the fact that relevant logic does not force the so-called fallacies of modality on us makes it a better vehicle than classical logic to act as a base for certain sorts of modality. Some philosophers claim that deontic modality is of this sort. Hume is supposed to have held that we cannot derive what ought to be the case merely from what is the case. Hume says:

In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in ordinary reasoning, and establishes the being of God, or makes some observation concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought*, or *ought not*, expresses some new relation or affirmation, 'tis necessary that it should be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from the others, which are entirely different from it. (*Treatise* Book II, Part I, section i)

Clearly, Hume is making a slightly weaker claim than is usually made in formulations of 'Hume's principle'. What he is claiming here is that anyone who thinks that we can derive an obligation from what is the case owes us an explanation of how this derivation can be made. What is usually called Hume's

principle, however, is the principle that we cannot derive any obligations from any set of non-moral facts.

Arthur Prior formulated what is now a famous objection to Hume's principle. He claimed that there are cases when we can infer obligations from pure descriptions. For example, the following inference appears to be valid (Prior 1976):

$\text{Fire Chiefs are municipal officials.}$
$\therefore \text{Fire Chiefs ought to do whatever all municipal officials}$ ought to do.

On the surface at least, it looks as though Hume's principle is violated by this argument. In addition, the argument seems valid. Which is correct, Hume's principle or the counterexample?

It turns out that Prior's counterexample, properly understood, does not violate Hume's principle as we have formulated it. We can formalise the Prior argument as

$$\frac{\forall x(Fx \rightarrow Mx)}{\therefore \forall P(\forall x(Mx \rightarrow OPx) \rightarrow \forall x(Fx \rightarrow OPx))}$$

Note that in this formalisation we have conflated doing something with being something. This may be a very problematic conflation in many cases, but I don't think it will do any harm here. More importantly, what we should note is that this inference is valid in relevant logic. It follows from basic principles of quantification (see below) and the transitivity of implication.

Having set Prior's argument aside, then, we can see that relevant deontic logic has a virtue that its classically based counterpart does not share. Relevant deontic logic does not force any violations of Hume's principle on us. Thus I suggest that relevant logic is a better contender for a basis for deontic logic than classical logic.

We will return to the topic of deontic logic in chapter 11 below. This concludes our main discussion of modality in this chapter, and we move on to the topics of quantification and identity.

6.10 Introducing quantification

The motivations that I have given for relevant logic have to do with capturing ordinary deductive inference and formalising some pieces of natural language (in particular, conditionals). When it comes to quantification, what philosophers and logicians have traditionally examined is not natural language quantification, but rather what we will call 'philosophers' quantification'. Quantifiers are phrases like 'all' and 'some', as well as 'most', 'many', 'there are at least two of',

and so on. Philosophers and logicians usually only deal with ‘all’ and ‘some’, and those other quantifiers that can be defined from those with the help of the other logical connectives. There is a branch of quantification theory known as the *theory of generalised quantifiers* that treats this other list as well. But we will stick to the standard quantifiers here – they will give us enough trouble. Natural language quantification is always restricted in a certain way. Consider the sentences,

Everyone is happy

and

Everything is working well.

If someone were to say the first of these sentences, you would not think that she meant that everyone in the universe is happy. Rather, you would look to features of context which told you to which group of people she was referring. Likewise, if someone were to utter the second sentence, you would not think that everything in the universe is working well, but rather that some collection of machines or whatever that are salient to the conversation are working well. Philosophers’ quantification, on the other hand, has as its domain everything in the universe, or at least everything of a given logical type. The quantified formula ‘ $\forall x F(x)$ ’ tells us that every individual in the universe is *F*. If ‘*F*’ means, say, red, then this formula means that every individual in the universe is red. Clearly, people ordinarily do not talk this way.

Situation semantics deals quite well with natural language quantification. Situation semanticists distinguish between a discourse situation (or context of utterance) and a resource situation. When one says, for example, ‘everyone is happy’, we look to features of the discourse situation to indicate what resource situation is meant. Then, to evaluate the statement, we look to the resource situation to see whether everyone in it is happy (see (Barwise and Perry 1983)). How to integrate the situated theory of quantification into our semantics for relevant logic is not a topic I have considered, nor has anyone else to my knowledge. It would seem an interesting avenue for future research.

Relevant logicians have, however, developed a theory of the philosophers’ quantifiers. Although philosophers’ quantification is not the same as ordinary language quantification, it is good to have a theory of it. The history of modern logic shows that philosophers’ quantification is an extremely useful technical device. It was used to formulate Russell’s theory of descriptions (one of the crowning achievements of modern logic) and has been used in the foundations of mathematics and for various uses in doing formal metaphysics. Thus, it would be good if relevant logic could be extended by adding philosophers’ quantifiers. It is to this form of quantified relevant logic that we now turn. But before we

can get into the theory of quantification proper, however, we need to understand a few facts about quantified logic.

Most readers of this book, I assume, have studied some quantification theory. It is a part of standard introductory courses in formal logic. But still some readers may need a bit of a refresher course, so here are some basic facts about quantification. We are adding the quantifiers $\forall$ (all) and $\exists$ (some) to the language. These bind variables (here we use letters from the later part of the alphabet as variables) in formulae. In the formula $\forall x F(xy)$, for example, the variable x is bound and the variable y is free. In the formula $\forall x F(x) \rightarrow G(x)$, the first occurrence of x is bound and the second occurrence of x is free. On the semantic side, quantifiers range over *domains*. For a quantified sentence, say, $\forall x F(x)$, we talk about whether an object from the domain satisfies or does not satisfy the open formula (a formula with perhaps some free variables) $F(x)$. An intuitive picture of satisfaction is not difficult to give. Suppose that we have a sentence $B(x)$, which means that x is blue. An object satisfies this sentence if and only if it is blue. Thus, when we establish a meaning for the predicates of our language, we can decide what things do or do not satisfy open formulae. With regard to situations, which objects satisfy which formulae is also easy, at least once the meanings of the predicates have been stipulated. For if the state of affairs $\langle \text{Blue}, a \rangle$ is in a situation s and the predicate B that means ‘blue’, then a satisfies $B(x)$ at s .

Now let us move to the treatment of the quantifiers in the natural deduction system for relevant logic.

6.11 Natural deduction rules for quantifiers

Natural deduction rules for quantification, both for classical and non-classical logics, can be quite tricky. The difficulties lie mostly in getting a sound formulation of the introduction rule for the universal quantifier and the elimination rule for the existential quantifier. The other rules are quite straightforward.

First, we have the elimination rule for the universal quantifier. This is what one would expect. It says that we can infer an instance from a universally quantified statement. Thus we have

$$(\forall E) \text{ From } \forall x A_\alpha \text{ to infer } A[t/x]_\alpha,$$

where ‘ $A[t/x]$ ’ means ‘the result of replacing all free occurrences of x with t such that t remains free’ (where t is an arbitrary term). We also have the dual rule for the existential quantifier, i.e. the rule that allows us to infer an existentially quantified statement from any of its instances, viz.,

$$(\exists I) \text{ From } A_\alpha \text{ to infer } \exists x A(x/t)_\alpha,$$

where ' $A(x/t)$ ' means 'the result of replacing zero or more occurrences of a term t with x such that x remains free in the resulting formula' (although of course x is bound in ' $\exists x A(x/t)$ '). Note that in the case of the existential quantifier introduction, we are allowed to generalise on any number of occurrences of a term in the statement.

Onto the more difficult cases. We begin with universal quantifier introduction ($\forall I$). In order to treat ($\forall I$), Anderson and Belnap borrow a device from Frederick Fitch. This device is the 'generalised categorical subproof' (see (Anderson *et al.* 1992), §31.1). It is perhaps easiest to introduce this idea using an example. Consider the following proof of $\forall x(F(x) \rightarrow G(x)) \rightarrow (\forall x F(x) \rightarrow \forall x G(x))$:

1.	$\forall x(F(x) \rightarrow G(x))_{\{1\}}$	<i>hyp.</i>
2.	$\forall x F(x)_{\{2\}}$	<i>hyp.</i>
3.	$\forall x F(x)_{\{2\}}$	2, <i>reit</i>
4.	$x \mid \forall x(F(x) \rightarrow G(x))_{\{1\}}$	1, <i>reit</i>
5.	$F(x) \rightarrow G(x)_{\{1\}}$	4, $\forall E$
6.	$F(x)_{\{2\}}$	3, $\forall E$
7.	$G(x)_{\{1,2\}}$	5, 6, $\rightarrow E$
8.	$\forall x G(x)_{\{1,2\}}$	3 – 7, $\forall I$
9.	$\forall x F(x) \rightarrow \forall x G(x)_{\{1\}}$	2 – 8, $\rightarrow I$
10.	$\forall x(F(x) \rightarrow G(x)) \rightarrow (\forall x F(x) \rightarrow \forall x G(x))_{\emptyset}$	1 – 9, $\rightarrow I$

Here we use the Fitch-style vertical line to indicate a categorical subproof. The variable to the left of the line indicates the variable on which we are going to generalise at the end of the subproof. It also indicates that we cannot reiterate a previous step into the subproof if that variable is free in the formula of the step. So, for example, we cannot reiterate $F(x)$ into a subproof if the variable indicated is x . Anderson and Belnap refer to the vertical line as a 'barrier', since it acts as a barrier to formulae with x free passing into the subproof.

Anderson and Belnap's existential quantifier elimination rule is constructed in analogy with the rule of disjunction elimination. Recall that disjunction elimination uses three premises: From $A \vee B_{\alpha}$, $A \rightarrow C_{\beta}$ and $B \rightarrow C_{\beta}$ to infer $C_{\alpha \cup \beta}$. In $\exists E$ we only need two premises:

($\exists E$) From $\exists x A_{\alpha}$ and $\forall x(A \rightarrow B)_{\beta}$ to infer $B_{\alpha \cup \beta}$, where x is not free in B .

As we said before, this rule is very much like disjunction elimination, and both are closely related to implication elimination. In the rule of existential quantifier elimination, the first premise tells us that there is something that is such that A and the second tells us that for anything, if it is A , then B . So we can

infer that B . The restriction that x not be free in B saves us from being able to infer non-theorems like $\exists x F(x) \rightarrow \forall x F(x)$. The close relationship between disjunction elimination and existential quantifier elimination is natural, because the existential quantifier works in some ways like an infinite disjunction. For, if we have a name for each element in our domain, $\exists x A$ tells us that $A[a/x]$ or $A[b/x]$ or $A[c/x]$

Let's look at a sample proof using this rule. This proof of $\forall x(F(x) \rightarrow G(x)) \rightarrow (\exists x F(x) \rightarrow \exists x G(x))$ is taken (with small changes) from Anderson and Belnap ((Anderson *et al.* 1992), p. 15):

1.	$\forall x(F(x) \rightarrow G(x))_{\{1\}}$	<i>hyp.</i>
2.	$\exists x F(x)_{\{2\}}$	<i>hyp.</i>
3.	$x \mid F(x)_{\{3\}}$	<i>hyp.</i>
4.	$x \mid \forall x(F(x) \rightarrow G(x))_{\{1\}}$	1, <i>reit</i>
5.	$x \mid F(x) \rightarrow G(x)_{\{1\}}$	4, $\forall I$
6.	$x \mid G(x)_{\{1,3\}}$	3, 5, $\rightarrow I$
7.	$x \mid \exists x G(x)_{\{1,3\}}$	6, $\exists I$
8.	$x \mid F(x) \rightarrow \exists x G(x)_{\{1\}}$	3 – 7, $\rightarrow I$
9.	$\forall x(F(x) \rightarrow \exists x G(x))_{\{1\}}$	3 – 8, $\forall I$
10.	$\exists x G(x)_{\{1,2\}}$	2, 9, $\exists E$
11.	$\exists x F(x) \rightarrow \exists x G(x)_{\{1\}}$	2 – 10, $\rightarrow I$
12.	$\forall x(F(x) \rightarrow G(x)) \rightarrow (\exists x F(x) \rightarrow \exists x G(x))_{\emptyset}$	1 – 11, $\rightarrow I$

Here we can see the interaction of the universal quantifier introduction rule and the existential quantifier elimination rule. The two are often used in conjunction, because the existential quantifier rule requires for its second premise a universally quantified formula.

There are some other 'house-keeping' rules needed to round out the treatment of quantification. One of these treats the interaction of quantification and disjunction and the other treats the interaction of quantification and conjunction. The first is called ' $\forall\vee$ ' by Anderson and Belnap, but sometimes goes by the name 'confinement', viz.,

From $\forall x(A \vee B)_{\alpha}$ to infer $A \vee \forall x B_{\alpha}$, where x is not free in A .

The second is called ' $\wedge\exists$ ' :

From $A \wedge \exists x B_{\alpha}$ to infer $\exists x(A \wedge B)_{\alpha}$, where x is not free in A .

Given these rules, we can prove many standard theorems about quantification.

For example, now, here is a proof of $\forall x(A \vee B) \rightarrow (\exists x A \vee \forall x B)$:

1.	$\forall x(A \vee B)_{\{1\}}$	<i>hyp</i>
2.	$x \mid \forall x(A \vee B)_{\{1\}}$	1, <i>reit.</i>
3.	$A \vee B_{\{1\}}$	2, $\forall E$
4.	$A_{\{2\}}$	<i>hyp.</i>
5.	$\exists x A_{\{2\}}$	4, $\exists I$
6.	$\exists x A \vee B_{\{2\}}$	5, $\vee I$
7.	$A \rightarrow (\exists x A \vee B)_{\emptyset}$	4 – 6, $\rightarrow I$
8.	$B_{\{3\}}$	<i>hyp</i>
9.	$\exists x A \vee B_{\{3\}}$	8, $\vee I$
10.	$B \rightarrow (\exists x A \vee B)_{\emptyset}$	8 – 9, $\rightarrow I$
11.	$\exists x A \vee B_{\{1\}}$	3, 7, 10, $\vee E$
12.	$\forall x(\exists x A \vee B)_{\{1\}}$	2 – 11, $\forall I$
13.	$\exists x A \vee \forall x B_{\{1\}}$	12, $\forall\vee$
14.	$\forall x(A \vee B) \rightarrow (\exists x A \vee \forall x B)_{\emptyset}$	1 – 13, $\rightarrow I$

We can also show that the theorems of this logic that just include conjunction, disjunction, negation, the quantifiers, variables, predicate letters and parentheses (i.e. that do not include implication), are exactly the theorems of classical first order logic. This means that all of classical first order logic can be captured in a relevant logic.

6.12 *Semantics for quantification

(This section is directed at technically sophisticated readers and very little that follows depends on it.)

A semantics for quantified relevant logics has been developed by Kit Fine. The semantics is ingenious and very complicated.⁵ Here we will not go into all the details of the semantics. I give those in an appendix at the end of the book. Rather, here I will try to give the reader the intuitions behind Fine's semantics.

Let us start with the standard treatment of quantification in modal logic. Let w be a possible world. Then, according to the standard semantics, the formula $\forall x A$ is true at w if and only if every individual in the domain of w satisfies the formula A . Translating this into the idiom of our semantics, it would seem that we get the following formulation. Let s be an arbitrary situation, then the formula $\forall x A$ is true at s if and only if every individual in the domain of s satisfies the formula A . This clearly is not what we want. As we have said, one of the virtues of using situations instead of worlds is that situations do not contain the entire

⁵ Bob Meyer and I have tried hard to simplify this semantics. We also had a series of discussions with Fine about this project in 1990 when all of us were at the Automated Reasoning Project at The Australian National University. Unfortunately, we did not succeed in producing a simplified semantics. I recommend this problem to the reader with the caution that it is very difficult indeed.

domain of the worlds that contain them. Situation theorists use this to explain how ordinary definite descriptions refer and how natural language quantifiers work. In order to give a situated treatment of philosophers' quantifiers what we need is a theory that explains how a situation nevertheless can represent a world as being such that everything in that world satisfies open sentences.

The basic idea that Fine uses to capture this idea is from Kripke's semantics for intuitionist logic. In the semantics for intuitionist logic situations have different domains and if $s \sqsubseteq s'$, for any situations s and s' , the domain of s is contained in the domain of s' . The truth condition for the universal quantifier says that a situation s makes the formula ' $\forall x A$ ' true if and only if for all situations s' such that $s \sqsubseteq s'$, every object in the domain of s' satisfies A at s' . In this semantics, if s makes true ' $\forall x A$ ', then s contains the information that, whatever objects there are they satisfy ' A '. Fine translates this idea into the semantics for relevant logic.

But this translation is not very straightforward. The problem has to do with the relationship between variable domains and the ternary accessibility relation. In particular, if we allow variable domains it becomes very difficult to use the standard semantic conditions to verify their correlated axiom schemes. In order to deal with this problem, Fine constructs his model from a set of models for propositional relevant logic. With each propositional model is associated a domain of individuals. If a formula ' A ' contains only individual constants that refer to items in the domain of a situation s , we say that ' A ' is in the vocabulary of s . The various situations in the different propositional models bear important relations to one another. We will use the letter Q to indicate the chief of these relations. For situations s and t , Qst only if the domain of s is a subset of the domain of t . The truth condition for the universal quantifier is

' $\forall x A$ ' is true at s if and only if for all t such that Qst ,
every individual in the domain of t satisfies ' A ' at t .

Fine places conditions on Q and on the other operators and relations of his theory so that he can prove the following lemma. *For any situation s and any formula A in the vocabulary of s , A is true at s if and only if for all situations t such that Qst , A is true at t .* Because of the truth of this lemma, we can call the situations that are related to s by Q , *conservative extensions of s* . That is, these other situations say the same things as s in the vocabulary of s .

To see why it is useful to have conservative extensions so closely linked with the truth condition for the universal quantifier, consider the following proof of the validity of ' $\forall x(A \vee B) \rightarrow (A \vee \forall x B)$ ', where x does not occur free in ' A '. We take an arbitrary situation s and suppose that ' $\forall x(A \vee B)$ ' is true at s . Then we assume that t is such that Qst . We now want to prove that ' $A \vee \forall x B$ ' is true at s . So we suppose that A does not obtain at s and show that ' $\forall x B$ ' does. Since ' $\forall x(A \vee B)$ ' is in the vocabulary of s , ' A ' is also in the vocabulary of t .

By the lemma stated above showing that t must be a conservative extension of s , if ' A ' fails in s it also fails in t . But, by assumption and the truth condition for the universal quantifier, ' $A \vee B$ ' is true in t . Thus, by the truth condition for disjunction, ' B ' is true in t . The situation t was chosen arbitrarily, and so we can generalise to say that ' $A \vee \forall x B$ ' is true at s . This suffices to prove that ' $\forall x(A \vee B) \rightarrow (A \vee \forall x B)$ ' is valid.

It turns out, as Fine indicates, that we can derive an intuitionist-like truth condition for the universal quantifier. Let us define a relation $\sqsubseteq$ such that $s \sqsubseteq t$ if and only if there is some situation s' such that Qss' and $s' \leq t$. Intuitively, a statement ' $s \sqsubseteq t$ ' says that t extends s perhaps in domain as well as in information. The truth condition now can be stated as

' $\forall x A$ ' is true at s if and only if for all t such that $s \sqsubseteq t$
every individual in the domain of t satisfies ' A ' at t .⁶

Thus, we can interpret a formula ' $\forall x A$ ' as saying that A is true of every individual in every extension of the current situation, or as saying that we have the information that ' A ' is true of every individual no matter how the current situation is extended in information or in domain.

Fine's technical achievement in producing this semantics is considerable. It is extremely subtle and fine-tuned. But we would like a simpler semantics for quantification and one that is easier to interpret philosophically. To provide such a semantics is one of the major open problems in relevant logic.

6.13 The semantics of identity

The standard semantics for identity sets ' $a = b$ ' true if and only if ' a ' and ' b ' refer to the same thing. This truth condition is, of course, very intuitive, but we cannot accept it except in a rather modified form. For consider the names 'Tully' and 'Catiline'. They do not refer to the same object. If the theory of rigid designation is correct, then a name refers to the same thing in every situation in which it refers at all. So, it would seem that in no situation do we have

cicero = catiline.

Thus, by our semantics for implication, if we accept the standard semantics for identity, we also have to accept that

cicero = catiline $\rightarrow A$,

for every proposition A . This, however, is an intolerable irrelevance.

We get around this problem by making identity into a standard binary relation, in the sense that we allow a pair (i, j) to be in the extension of identity at a

⁶ This is, in effect, the 'orthodox' truth condition for the universal quantifier stated in Fine (1988).

situation where i and j are in fact (i.e., from the standpoint of our metalanguage) distinct individuals. We do not merely do this in order to solve the problem cited above. For we have uses for situations that identify objects that are in fact distinct. For example, there are fictions in which things that are really distinct are said to be the same. There is an episode of the science fiction television programme, *Star Trek*, in which a very old man is supposed to have been various people in history, such as Caesar and Brahms. In order to model such fictions, it is useful to have situations that identify really distinct objects (Mares 1992).

In order for standard inferences about identity to follow in our semantics, we have to place certain constraints on the extension of identity. Identity is transitive, in the sense given in the following deduction:

1. $a = b \wedge b = c_{\{1\}}$ *hyp*
2. $a = c_{\{1\}}$ 1, *trans* =

As we saw in chapter 5 above, we reject the ‘nested’ form of transitivity, but this conjunctive form is appropriate for relevant logic. We make the conjunctive form of transitivity valid by simply postulating that if (i, j) is in the extension of identity at a situation s , and (j, k) is also in the extension of identity at s , then so is (i, k) . Similarly, in order to satisfy the symmetry of identity (which allows the inference from $a = b_\alpha$ to $b = a_\alpha$) we postulate that, if (i, j) is in the extension of identity at a situation, then (j, i) is also in that extension at that situation.

Traditionally, there are four features that distinguish identity – symmetry, transitivity, reflexivity, and the principle of the substitution of identicals. Reflexivity is treated in a very straightforward manner. A relation φ is reflexive if and only if, for every entity a , φaa . For identity, we postulate that, for every logical situation, s , is such that for everything, i , in the domain of that s , (i, i) is in the extension of identity.

The semantic treatment of the principle of the substitution of identicals is a little more involved. One standard version of the principle is the following:

$$a = b \rightarrow (A \rightarrow A(b/a)).$$

This is not a principle that a relevant logician could accept. For consider a formula A in which a does not occur. Then we have as an instance of the above principle,

$$a = b \rightarrow (A \rightarrow A).$$

Clearly, this is irrelevant. So, the first change we make is to weaken this principle to

$$(a = b \wedge A) \rightarrow A(b/a).$$

This weaker principle is called ‘weak substitution’. Now if A does not contain a , we get the harmless thesis,

$$(a = b \wedge A) \rightarrow A.$$

But there still are difficulties with weak substitution. One problem is of a technical nature and that has to do with Dunn’s theory of relevant predication.⁷ But there are relatively non-technical reasons as well why we should abandon weak substitution.

Accepting weak substitution leads to a watering down of the relevance properties of implication. Let m stand for ‘my favourite flavour of ice cream’, c mean ‘chocolate ice cream’, and s mean ‘strawberry ice cream’. Then we have

$$(c = m \wedge (s = m \rightarrow s = m)) \rightarrow (s = m \rightarrow s = c).$$

This statement has a true antecedent. ‘Chocolate is my favourite flavour and strawberry is my favourite flavour’ does imply itself (at least in worldly situations). But it is not the case that ‘Strawberry is my favourite flavour’ implies ‘Strawberry ice cream is identical to chocolate ice cream’, if implication is taken to be the sort of robust relevant connective that we have been discussing. So, we have an implication with a true antecedent and a false conclusion, and so we must reject weak substitution.

One problem with both strong and weak substitution is that they allow us to substitute for identicals into statements that contain implications. Implication is an intensional connective, and we should expect that it would give us problems with substitution. So here is an even weaker principle of substitution that would seem harmless:

$$(a = b \wedge A) \rightarrow A(b/a),$$

with the proviso that A be implication free. So far I have not heard any objections to this principle, so I propose that we accept it at least for now.

6.14 Identity and proper names

Here is a further problem. Consider two names that refer to the same individual, such as ‘Cicero’ and ‘Tully’. On the semantics that we have put forward, the following inference is valid in our intended model:

- | | |
|---------------------------------------|---------------------------------|
| 1. $cicero = a_{\{1\}}$ | <i>hyp</i> |
| 2. $tully = a_{\{1\}}$ | 1, <i>truth condition for</i> = |
| 3. $cicero = a \rightarrow tully = a$ | 1 – 2, $\rightarrow I$ |

⁷ For the theory of relevant predication, see chapter 11 below. For the problem, see Mares 1992.

The conclusion is one that a relevant logician should find a little unpleasant. But things get worse. The following argument is also valid in our intended model:

- | | |
|--|---------------------------------|
| 1. $cicero = cicero_{\{1\}}$ | <i>hyp</i> |
| 2. $tully = cicero_{\{1\}}$ | 1, <i>truth condition for</i> = |
| 3. $tully = tully_{\{1\}}$ | 2, <i>truth condition for</i> = |
| 4. $cicero = cicero \rightarrow tully = tully_{\emptyset}$ | 1 – 3, $\rightarrow I$ |

This conclusion seems quite irrelevant.

Some philosophers, however, might tell us to accept this conclusion. A direct reference theorist, for example, might tell us that ' $cicero = cicero \rightarrow tully = tully$ ' is quite acceptable. On a standard version of the theory of direct reference, sentences express structured propositions. For example, the sentence 'Ramsey is a dog' expresses the structured proposition $(Dog, ramsey)$, where '*Dog*' refers to the property of being a dog and '*ramsey*' refers to my dog. On the theory of direct reference, a proper name contributes only its referent to the structured proposition expressed by the sentence. This is to be contrasted with a definite description which contributes the properties used to describe the object. Since the referent of '*Cicero*' and '*Tully*' are the same, the structured proposition expressed ' $cicero = cicero$ ' is $(=, cicero, cicero)$ and the proposition expressed by ' $tully = tully$ ' is $(=, cicero, cicero)$. Thus, a direct reference theorist could well say that the implication at step 4 of the above proof is quite relevant, that is, it is merely a self-implication.

On the other hand, we can modify our semantics slightly in a way that is in accord with the theory of direct reference that avoids commitment to these sorts of seemingly irrelevant implications. On Kaplan's theory of direct reference, in addition to its content, an expression has a *character*. To understand what a character is, we first have to understand the distinction between contexts of utterance and contexts of evaluation. A context of utterance is merely a context (in our theory, a situation) in which an expression is said, written, signed, or whatever. The context of utterance determines the referents of the expressions used in an utterance. For example, I yelled 'Ramsey stop barking!' just before writing this sentence. My use of 'Ramsey' refers to my dog. Suppose at the very same time the philosopher Hugh Mellor in Cambridge said 'Ramsey was a great philosopher'. He is not referring to my dog, but to a human being, Frank Ramsey. The context of evaluation, on the other hand, is a situation in which we determine whether a statement is true or false. Suppose for example, that I say 'Ramsey could have been an obedience champion' we first use the context of utterance to determine that the referent of 'Ramsey' is my dog. In so doing we determine the structured proposition $(Obedience\ Champion, ramsey)$. We then evaluate this proposition at all situations accessible to the present one to

determine whether any make it true. These are contexts of evaluation. If at least one context of evaluation does make it true, then the statement that I made is correct.

The character of an expression is a function from contexts of utterance to its content. For a proper name, its character is a function from contexts of utterance to its referents. For example, the character of 'Ramsey' is a function that takes my current situation to my dog and takes Hugh Mellor's current situation to Frank Ramsey.

We can use the notion of a character to modify our semantics. Instead of having states of affairs (SOA) that contain the referents of expressions (i.e. individuals and properties), we could have SOA that contain the characters of expressions. Since the character of 'Cicero' is distinct from the character of 'Tully', we could include the 'SOA' $\langle =', cicero', cicero' \rangle$ in a situation s without including $\langle =', tully', tully' \rangle$ in s , where the primes indicate characters rather than properties or individuals.

Although I do not wish to commit myself to the theory of direct reference, it is nice to know that we can amend our semantics both to be consistent with this theory and to avoid grating implications like $cicero = cicero \rightarrow tully = tully$.

6.15 Necessity and identity

The preceding discussion raises some interesting issues regarding the relationship between identity and modality in our framework. As is well known, Kripke (and others following him) argues that all true identity statements are necessarily true (Kripke 1972). To make Kripke's argument run, we need to accept the doctrine of direct reference and the view that objects can exist in more than one possible world. Suppose that ' a ' and ' b ' are proper names. If we accept the theory of direct reference, then they have no meaning except in the sense of a character, as explained above, and merely refer to their referent. They do not refer by means of expressing a meaning. Let's fix a context of utterance. In this context of utterance, the name ' a ', say, refers to the object i . If ' a ' has a referent in a given possible world, then that referent is i . Thus, ' a ' is said to be a *rigid designator* – it refers to the same object in every world in which that object exists. Now consider both ' a ' and ' b '. If they refer to the same object, given our context of utterance, then they refer to that object in every possible world (taken as a context of evaluation). Since they are rigid designators, they cannot refer to the same thing in one world and to different things in another world.

Now, this is a very brief and unsatisfactory rendition of Kripke's argument. But that does not matter here. What does matter to us is the consequence that identity statements are necessary. There are two ways of interpreting

this position. The first, is to incorporate what we call ‘Kripke’s principle’, viz.,

$$a = b \rightarrow \Box a = b \text{ (Kripke’s Principle).}$$

This is clearly very easy to incorporate into our semantics. Given a situation s , if s' is modally accessible from s , then the true identities at s' include all of those that hold at s . But the question is whether we really should make valid Kripke’s principle.

It would seem that we can accept Kripke’s argument without buying into Kripke’s principle. We can make the true identities at every worldly situation agree with one another. This will make every identity necessary in the sense of being necessary in the model. But this does not mean that identity statements imply their own necessity. It would seem that we need a further argument for that view.

For more about relevant logic and identity, I recommend Philip Kremer’s paper (Kremer 1999).

6.16 A useful theorem

It will be useful for the theory of conditionals developed in part II of this book to prove that, for normal modalities, only worldly situations are accessible from worldly situations.

In order to prove this, we will need the following postulates:

1. If Mst , Ms^*t^* .
2. If $s \sqsubseteq s'$ and $Ms't'$, then there is some $t \sqsubseteq t'$ such that Mst .
3. If $s \sqsubseteq t$, then $t^* \sqsubseteq s^*$.

We now prove the following lemmas:

Lemma 7 *If $s \sqsubseteq s'$ and Mst , then there is some situation t' such that $t \sqsubseteq t'$ and $Ms't'$.*

Proof. Suppose that $s \sqsubseteq s'$ and Mst . By postulate 1, Ms^*t^* . By postulate 3, we also know that $s'^* \sqsubseteq s^*$. By postulate 2, then, we obtain that there is some $t'^* \sqsubseteq t^*$ such that $Ms'^*t'^*$. By postulate 1, and the fact that $s'^{**} = s'$ and $t'^{**} = t'$, $Ms't'$. We also know that $t \sqsubseteq t'$. Thus, we have proven the lemma. (QED)

Lemma 8 *If s and s^* are both logical situations, then $s = s^*$.*

Proof. Suppose that s and s^* are both logical situations. Thus it is both the case that $s \sqsubseteq s^*$ and $s^* \sqsubseteq s$. Since $\sqsubseteq$ is antisymmetrical, $s = s^*$. (QED)

These lemmas allow us to derive the following theorem:

Theorem 9 *If M is an accessibility relation for a normal modal logic, then if $s \in Poss$ and Mst , t is also in $Poss$.*

Proof. Suppose that M is the accessibility relation for a normal modal logic and that s is a possible situation. Also assume that Mst . By the definition of $Poss$, there is some worldly situation s' such that $s \leq s'$. Then, by lemma 7, there is some t' such that $t \leq t'$ and $Ms't'$. By postulate 1, $Ms'^*t'^*$. Since s' is worldly, $s^* = s$. And, since M is normal, t' is logical. By postulate 1, $Ms'^*t'^*$, i.e., $Ms't'^*$. Since M is normal, t'^* is logical as well. So, by lemma 9, $t' = t'^*$. Thus, t' is worldly. So t is possible, ending the proof of the theorem. (QED)

6.17 Summary of part I

We have now come to the end of Part I of the book. Before we go on it will be useful to summarise what we have done so far. First, we constructed an ontology of situations. We used this ontology to support an interpretation of Israel and Perry's theory of information and, in turn, to develop an informational interpretation of relevant implication. We then extended this interpretation to include both negation and modality.

The aim of this exercise was to provide an understanding of the connectives of propositional relevant logic. We also wanted to give an interpretation in terms that many philosophers of logic can live with. We used a reasonably conservative ontology of worlds, individuals, and (non-well-founded) sets. Thus we tried to defend relevant logic against two objections. The first of these objections is that the semantics of relevant logic is merely a mathematical theory that does not yield a real theory of meaning for the connectives. Throughout this part of the book, we have argued that the semantics of relevant logic, properly understood, gives the connectives intelligible and intuitive meanings. The second objection is that the semantical theory for relevant logic is committed to strange entities such as impossible situations. The reductionist project of chapter 4 is meant to show that the commitments of this semantical theory are rather more harmless than the objection indicates.

In Part II, we will use the Routley–Meyer frame-theory, and our interpretation of it, to formulate a semantics for indicative conditionals and a semantics for counterfactuals.

Part II

Conditionals

7 Indicative conditionals

7.1 What are indicative conditionals?

One way in which I motivated the acceptance of relevant logic in the introductory chapter above is as a basis for a theory of conditionals. The vast majority of the writers on conditionals distinguish between two sorts of conditionals – indicative conditionals and counterfactual conditionals. I will follow this tradition here. We begin with the indicative conditional.

The best way I know to distinguish between indicative conditionals and counterfactual conditionals is to cite some examples. And the best examples I know are these old chestnuts due to Ernest Adams:

- If Oswald didn't kill Kennedy, someone else did. (7.1)
If Oswald hadn't killed Kennedy, someone else would have. (7.2)

The first of these – the indicative conditional – assumes for its utterance that Kennedy was killed, but not that Oswald killed him. The second – the counterfactual – assumes that Kennedy was killed and that Oswald did it.

But this isn't the only difference between the two. The indicative conditional seems to express something about the world and something about what we believe. For consider the following conditional:

- If Oswald didn't kill Kennedy, there is a cover up of the fact (7.3)
that Kennedy is still alive.

Even if we have a firm belief that Oswald did murder Kennedy, those of us who are not conspiracy theorists will be more prepared to accept (7.1) than (7.3). The reason for this is, in part, that we believe rather strongly that Kennedy was killed and this belief does not depend on our belief that Oswald killed Kennedy, i.e., even if we abandon the latter belief we will not abandon the former, nor will our degree of belief drop in the proposition that Kennedy is dead. This indicates that our reasons for asserting conditionals at least to some extent has to do with connections between our beliefs.

As we shall see, proponents of Adams' thesis hold this view very strongly, but my view also accords with this connection between beliefs and conditionals to some extent.

Whereas the role of the indicative conditional is at least partly doxastic, the role of the counterfactual conditional is more ontological. As we shall see in chapter 8 below, counterfactuals are used to express dispositional facts about individuals, populations and structures of other sorts. They have other roles too, but for the most part they are used to talk about the nature of things. Yet, as we shall see in chapter 8 below, there is a much stronger connection between counterfactuals and indicative conditionals than is ordinarily supposed and there is a doxastic component in the semantics of counterfactuals.

There has been a rather active debate about which sorts of natural language conditionals should be analysed as counterfactuals and which should be characterised as indicative conditionals. In what follows, I will go against the more popular view and treat all and only subjunctive conditionals as counterfactuals. In particular, this means that I treat future tense indicative conditionals formally as indicative conditionals. Nothing deep or crucial rests on this point. All the examples in which I use future tense indicatives could be recast using past or present tense conditionals, although sometimes slightly more awkwardly.

In what follows, I set out a theory the grandparent of which is my and André Fuhrmann's theory of relevant conditionals (Mares and Fuhrmann 1995).

7.2 **Conditionals and relevance**

As I argued in chapter 1 above, conditionals express a connection of relevance between antecedents and consequents. Let's look briefly again at my argument.

What is needed is a theory of conditionals which demands some real relation between the antecedent and consequent of a true conditional. If the conditional is just material implication, then having a false antecedent alone is sufficient to make a conditional true. Thus, equating the conditional with material implication forces us to accept conditionals like this one:

If I pick up this guinea pig by its tail, its eyes will fall out.

Guinea pigs do not have tails, and so one cannot pick one up by the tail. But an utterance of this conditional indicates a real connection between picking up a guinea pig and its eyes' falling out. It seems to indicate more than just a relationship between the actual truth values of the antecedent and consequent.

Similarly, a true consequent is insufficient to guarantee the truth of a conditional. For consider an indicative version of Dunn's sentence,

If I scare this pregnant guinea pig, her babies will be born blind.

Again, the conditional indicates a stronger connection between antecedent and consequence than merely that the consequence is true regardless of the truth of the antecedent. The fact that guinea pigs are always born without sight is not enough to justify accepting the conditional.

In this chapter I develop a theory of the indicative conditional that satisfies the demand that there be a real relation between antecedent and consequent in true conditionals.

7.3 Indicative conditionals are not implications

In previous chapters, we have developed a theory of implication. As I said in the introductory chapter, I base a theory of conditionals on this theory of implication. I have been careful (I hope!) throughout the first few chapters not to say that the theory of implication itself provides us with a theory of conditionals. This is because, although they are closely related, implications are not themselves ordinary everyday conditionals.

The problem with implication, in regard to a theory of conditionals, is that both do too much and too little. There are rules that are permissible for implications that are not permissible for conditionals and there are rules that are permissible for conditionals that are not permissible for implications.

There are several other problems with identifying the indicative conditional with relevant implication. For example, there is a problem with the principle of contraposition. In the case that we will review, the problematic inference is from,

$$A \rightarrow \sim B,$$

to,

$$B \rightarrow \sim A.$$

Our treatment of implication and negation makes this inference valid. But when put in terms of the conditional, it becomes very counterintuitive. Consider the following argument due to Frank Jackson:

$$\begin{array}{l} \text{If Sally makes a mistake, she won't make a big one.} \\ \hline \therefore \text{If Sally makes a big mistake, she won't make a mistake.} \end{array}$$

Clearly the conclusion is absurd, even if the premise is true.

Rejecting contraposition as a principle governing implication is not an attractive solution. It is an intuitive principle. We use contraposition in proofs all the time. In logic and mathematics the contrapositive proof is one of the most commonly used proof techniques. Thus, dropping it gives us a radically counterintuitive theory of implication.

Problems also arise in the connection between implication and conjunction. On our view the following inference rule is valid:

From $A \rightarrow C_\alpha$, infer $(A \wedge B) \rightarrow C_\alpha$.

If we interpret the conditional to mean relevant implication, then, we also make valid the following argument:

If Ramsey was good today, Jane is happy.

∴ If Ramsey is good today and Jane has been given the sack,
Jane is happy.

Clearly, we do not want to accept this.

The above counterexample to strengthening of the antecedent, as this inference is called, is also a counterexample to the transitivity of the conditional. Anderson and Belnap's natural deduction system makes valid the following inference:

From $A \rightarrow B_\alpha$ and $B \rightarrow C_\beta$ infer $A \rightarrow C_{\alpha \cup \beta}$.

Moreover, it is easy to prove in that natural deduction system that $(A \wedge B) \rightarrow A$. So we have the following inference:

- | | |
|---|------------------------|
| 1. $A \rightarrow C_{\{1\}}$ | <i>hyp.</i> |
| 2. $A \wedge B_{\{2\}}$ | <i>hyp.</i> |
| 3. $(A \wedge B) \rightarrow A_\emptyset$ | <i>theorem</i> |
| 4. $A_{\{2\}}$ | 2, 3, $\rightarrow E$ |
| 5. $C_{\{1,2\}}$ | 2, 4, $\rightarrow E$ |
| 6. $(A \wedge B) \rightarrow C_{\{1\}}$ | 2 – 5, $\rightarrow I$ |

Thus, if we allow transitivity in this form, we also allow strengthening of the antecedent.¹ Thus, the fact that implication is transitive gives us another reason not to identify the indicative conditional with relevant implication.

7.4 But the indicative conditional is closely related to implication

We saw in the previous section that the indicative conditional is not relevant implication. But they do have a close relationship to one another. We saw in the introductory section to this chapter that the conditional, like implication, expresses some form of relevance between an antecedent and a consequent.

Moreover, an implication entails the corresponding conditional. For example, from the implication,

Two bodies' being positively charged will imply that they exert a repelling force on one another.

¹ Note that in this proof we bring in a theorem without proving it. We could have added the proof of the theorem to this proof. Since the subscript of a theorem is the empty set, we can use it very easily in implication eliminations.

we can infer the conditional,

If two bodies are positively charged, they will repel one another.

It seems that we want to make this sort of entailment generally valid. That is, if the formula,

$$A \rightarrow B,$$

holds at a situation, then the following formula should also be true there:

$$A \Rightarrow B,$$

where ' $\Rightarrow$ ' is the indicative conditional.

The story so far. The indicative conditional expresses a *weaker* connection between antecedent and consequent than relevant implication. But the conditional expresses a relevant connection of some sort.

I suggest that we can capture both of these features of the conditional by giving the conditional a modified form of the semantics for implication. Recall the Routley–Meyer ternary relation used to give a truth condition for implication.

In chapter 3 above, we interpreted the Routley–Meyer ternary relation on situations in terms of situated inference. Given the assumptions that the situation s is in a particular world and that the situation t is also in that world, we are licensed to infer that some situation is such that $Rstu$ is also in that world (and that the set of situations is such that $Rstu$ is the smallest set for which we have this licence).

I suggest that we assign to the conditional in a more restricted relation than Routley and Meyer's R . This means that, where C is the relation associated with the conditional, whenever there are three situations s , t , and u such that $Cstu$, then $Rstu$ holds as well. (In fact our relation, C , will be a bit more complicated than this. But we will leave those complexities for later.)

It might seem paradoxical to have a more restricted relation to give the semantics for a weaker connection between antecedent and consequent, but it really does make sense. For more propositions stand in a weaker relation than in a stronger relation. If we have fewer situations standing in an accessibility relation to one another, more propositions will stand to one another in the relation that this accessibility relation models. Thus, we need a stronger accessibility relation.

What we do is start with the set of situations that are related to our situation by the Routley–Meyer accessibility relation and then narrow our search space to pairs situations which are legitimately to be considered with regard to a given conditional. All this sounds rather complex, but we will break the idea down into manageable parts. To make things a bit easier, instead of the phrase, 'pairs of situations (t, u) such that $Rstu$ ', I use the word 'circumstances'.

One feature of ordinary discourse that will help to determine which situations are legitimately to be considered are the circumstances that the participants in

the conversation are willing to consider. This might sound tautologous, but it isn't. As we shall see, there are both objective and subjective factors that come into play in deciding which situations are legitimate.

Suppose that you say the following to me:

If Tim moves to Washington we won't see him again in Wellington this year.

And let's say that I disagree. But how can I rebut your conditional? It makes sense for me to retort with

Tim could get a well-paid job and fly home later this year.

But it would be useless for me to counter with

Tim might sprout wings and fly home whenever he wants.

In these replies to the original conditional, I am trying to point out circumstances in which it does not follow that Tim won't be in Wellington again this year, even if he does move to Washington. Some of these circumstances are reasonable to consider and others are ruled out of bounds. We would be willing to consider the circumstances in which Tim gets a good paying job, since he is a bright and personable fellow. But we would not be willing to consider circumstances in which he sprouts wings, since this is impossible.²

7.5 The role of the antecedent

If we are willing to take a conditional seriously, we must be willing to consider circumstances in which the antecedent is true. Consider again my favourite conditional,

If Sally has squared the circle you have lost your bet.

In a conversation in which this conditional is taken seriously, the participants of the conversation must include in the set of circumstances they are willing to consider, circumstances in which Sally squares the circle.

To make this more precise, if S is the set of circumstances used to evaluate ' $A \Rightarrow B$ ' at s , then there must be a pair (t, u) in S such that ' A ' is true at t , if there is any situation in our model in which ' A ' is true. In putting this condition on the set of considered circumstances, we are attempting to bar vacuously true conditionals as much as possible.

² One might think that the use of 'might' in 'Tim might sprout wings ...' just makes that statement false, since Tim's growing wings is impossible. But recall what we said in chapter 5 above: modalities in the relevant semantics do not all have to be treated as submodalities of metaphysical possibility. At any rate, it isn't clear that a human's growing wings is metaphysically impossible.

7.6 The role of the consequent

The consequent also plays an important role in determining the class of circumstances to be considered when evaluating a conditional. Let's think again about Dunn's conditional. Suppose that there is a pregnant guinea pig present. Someone says:

If I scare this guinea pig her babies will be born blind.

It may be that, like all baby guinea pigs, her babies will be born blind. But it would seem illegitimate only to consider circumstances in which her babies are born blind. For the evaluation of the conditional requires that we test the relationship between the antecedent and consequent of the conditional. And, for such a test, we need to see whether when the consequent fails to be true, the antecedent is absent as well.

Let's make this idea more precise. If S is the set of circumstances used to evaluate $A \Rightarrow B$ in s , then there is some circumstance (t, u) in S such that B fails to hold at u , if there are any situations in our model at which B fails to hold.

I used to think that there were exceptions to this rule. In particular, I used to accept the rule

$$\frac{A \Rightarrow B \quad \vdash B \rightarrow C}{\therefore A \Rightarrow C}.$$

Suppose that we are legitimately considering only circumstances in which C holds (in the second situation in each circumstance) in order to evaluate $A \Rightarrow B$. Suppose also that this conditional turns out to be true and that it is a theorem of our logic that $B \rightarrow C$. It would seem that we want to infer that $A \Rightarrow C$.

But I am no longer certain that we do want to make such inferences. In section 7.11 below, we will consider a set of circumstances in which 'Either Prof. Plum committed the murder or Col. Mustard did' are true (in the second situation in each circumstance). And we will argue that the conditional,

If Prof. Plum didn't commit the murder then Col. Mustard did.

is true, given that set. It is a theorem of our logic that 'Col. Mustard committed the murder' implies 'Either Prof. Plum committed the murder or Col. Mustard did.' Given the rule above, we can infer

If Prof. Plum didn't commit the murder either he or Col. Mustard did.

This sounds weird to me. I'm not sure that we want to accept it. If we reject it, we have to reject the rule above too.

The idea that the antecedent helps to determine the indices used to evaluate a conditional is commonplace in the philosophical literature. But the idea that the

consequent also plays a role is less common. But it has been proposed before, most prominently by Dov Gabbay (Gabbay 1972). We will look at Gabbay's motivation for his position in chapter 8 below.

7.7 Closeness

We pick the circumstances that we use to evaluate a conditional as the set of such that seems reasonable. But what counts as reasonable? When we introduced this notion above, we did so by looking at the sorts of retorts that we would take into account and those which we would ignore. Let's do that again.

Consider again our conditional about Tim:

If Tim moves to Washington we won't see him again for months.

In most contexts of utterance, it would not be legitimate to reply:

We will see him again if he proves that redness is a shape (and not a colour), goes on the chat show circuit and makes enough money to fly back and forth regularly.

This would be a ridiculous reply. Because the suggestion is metaphysically impossible, we would not be willing to consider it.

It would also be illegitimate as we have said, but slightly less so, to suggest that Tim could grow wings. This suggestion is less ridiculous because it is more plausible that Tim can grow wings than that he can prove redness to be a shape rather than a colour.

Perhaps still illegitimate but much less so is the following suggestion:

Well, we might see him if he wins the lottery and has the money to fly back to Wellington.

Here we have a suggestion that is not impossible but very implausible. We would take it more seriously than the others, but we might still treat it as an illegitimate suggestion in the context.

This ranking in terms of implausibility tells us something interesting about how we understand conditionals. We are willing to consider more plausible circumstances over less plausible ones. We rank metaphysically impossible circumstances as more implausible than nomically implausible ones (I take it that a person's growing wings is nomically implausible). And we rank nomically impossible circumstances as more implausible than remote possible, but nomically possible circumstances.

Thus, I think we can set out a hierarchy of modalities. *All things being equal*, if a circumstance is metaphysically impossible, we are more reluctant to count it among the ones that we consider to evaluate a conditional than ones that are metaphysically possible and if a circumstance is nomically impossible we are more reluctant to use it than one that is both nomically and metaphysically impossible.

Once again, it is time to be more precise and technical. We need to know exactly what it means for a circumstance to be possible. Here is a definition:

Let N be a type of modality. Then a circumstance (t, u) is possible relative to s in modality N if and only if there is some situation s' such that $M_N ss'$, $t \trianglelefteq s'$, and $u \trianglelefteq s'$.

This definition says that a circumstance (t, u) is possible relative to s if and only if both t and u are parts of a situation that is possible relative to s .

7.8 Putting the theory together

We have various parts of a theory of conditionals and we now need to assemble those parts into a coherent theory. In order to make the theory coherent, what is needed is a rule regarding which elements should take priority over others.

Here is such a rule.

A legitimate set of circumstances used to evaluate a conditional is a set of reasonably close circumstances that obeys the constraints regarding the antecedent and consequent.

This means that, if possible, there are circumstances in the set that satisfy the antecedent and those in which the consequent fails. The antecedent and consequent constraints take precedence over closeness. The idea is simple. We start with close circumstances and keep letting circumstances into our set until these conditions are fulfilled.

7.9 Modus ponens

So far in our treatment of conditionals we have been using our subjective inclinations as a guide to the structure of the set of circumstances that we choose. But it would seem that there are objective constraints on this set.

Consider a case in which a person says

If I flip this switch you will all die!

The villain then flips the switch, a funny whirring noise is heard, and then nothing happens. I would say that the villain was wrong when he made his original statement. We expect modus ponens to hold of conditionals. If we have a conditional with a true antecedent and a false consequent, we declare the conditional to be false.

But some philosophers disagree with this. There have been two sorts of counterexamples to modus ponens proposed in the literature. The first type uses nested conditionals, and originates with Vann McGee's paper (McGee 1985). McGee presents the following argument with (supposedly) true premises and a false conclusion. Imagine that the following is presented on the eve of the 1980

presidential race in the United States:

If a Republican does not win, then
 if Reagan does not win Anderson will win.
 A Republican will win

 $\therefore$ If Reagan does not win Anderson will win.

McGee is right in his claim that no one at all knowledgeable would assent to the conclusion of this argument. Anderson was a poor last in the polls. If Reagan had not won, surely Carter would have won. But the two premises look true. So what is wrong? The only inference rule that seems to have been used here is modus ponens. The conclusion that McGee draws is that modus ponens is not a valid inference rule.

Isaac Levi gives a similar counterexample ((Levi 1996), p. 106). We start with a wheel divided into three equal parts, 1, 2, and 3 and a stationary marker which indicates the winning position after the wheel has been spun. The wheel is spun and it lands on 1. Thus, we have the argument

If the winning segment is odd,
 if it does not land on 1, it will land on 3.
 The winning segment will be odd.

 $\therefore$ If the wheel does not land on 1, it will land on 3.

Here we have an inference of the same form as McGee's counterexample. Without the proviso that the wheel lands on an odd numbered segment, there is no reason to think that if it does not land on segment 1, it will land on segment 3. Thus, the conclusion is false, even though the premises are true.

Thus, the problem seems to be that we have an invalid argument of the form

$$\frac{A \Rightarrow (B \Rightarrow C) \quad A}{\therefore B \Rightarrow C}.$$

But, it would seem that the arguments above are not really of this form. The sticking point is the structure of the major premise. Consider again McGee's premise:

If a Republican wins, then if Reagan does not win Anderson will win.

Let's think about the set of circumstances one would use to evaluate this conditional. Clearly, one would look at reasonably close (and possible) circumstances in which Anderson and Reagan are Republicans, Carter is a Democrat, Reagan leads in the polls, Carter is second, and Anderson is well behind, and so on. We would not look at weird circumstances in which there have been wild sampling errors on the part of the pollsters, or any other circumstances that seem way out of line to us.

Using such circumstances, we can conclude that

If a Republican wins *and* Reagan does not win, Anderson will win.

But this is not a nested conditional – it is not of the form $A \Rightarrow (B \Rightarrow C)$. It does not seem that we can, in a natural way, constrain our set of circumstances to make the nested conditional true.

The nested conditional, however, sounds true. I suggest that the surface grammar of English is misleading in this way. Sometimes we are allowed to say a sentence that looks like a nested conditional but has as its logical form a conditional with a conjunctive antecedent (like the one given above). This is not a silly feature of English but rather has a useful function. Consider the difference between:

If Tim and Joe both have cars then Tim has a car. (7.4)

and

If Joe has a car and Tim has a money for petrol they will be able to get to Auckland. (7.5)

Both (7.4) and (7.5) are true. We would not be tempted, I think, to express (7.4) using (7.6):

If Tim has a car, then if Joe has a car Tim has a car. (7.6)

The sentence (7.6) seems false. But (7.7) below looks true:

If Joe has a car, then if Tim has money for petrol they will be able to get to Auckland. (7.7)

What is going on with (7.5) and (7.7) that is not happening in (7.4) and (7.6) is that both conjuncts of the antecedent are doing some work to get us to the consequent.

I think there is a reasonable informal notion of *doing some work* at play here but we can introduce a formal notion as well. Suppose that S is a set of circumstances attributed to ' $(A \wedge B) \Rightarrow C$ ' in s . And suppose that conditional is true in s . Then we say that A is doing some work in producing C if and only if there is some (t, u) in S such that B is true in t but C fails to be true in u . Thus, ' $(A \wedge B) \Rightarrow C$ ' can be expressed by 'If A , then if B , C ' in English if both A and B are doing some work in $(A \wedge B) \Rightarrow C$.³

This I think does away with the counterexamples to modus ponens that use nested conditionals. There are some, however, that use non-nested conditionals.

³ This is a sufficient condition. I am not sure that this is the only type of case in which we are allowed to express a non-nested conditional as a nested one.

The following is from Bill Lycan's (Lycan 1993) but he attributes it to Allan Gibbard:

If you insult me, I will not become violent.
If you insult my dog, I will become violent.
You insult both me and my dog.

Contradiction!

There is something very different going on here than in the previous examples. Suppose that I assert both of the conditional premises. And then you insult both me and my dog. Also suppose that after I react (or fail to react), some innocent and unharmed bystander asks me whether I still think my conditionals were true. I would say 'no'. If I have pummelled you I would reject the first premise and if I have not I would reject the second premise. Note that I am not merely saying that-would reject one or other of these premises as being false after the fact, but that I would say after the fact that they hadn't been true before your insult.

Thus, my reply to the Gibbard-Lycan example is to say that in any possible situation in which the third premise is true one or other of the conditional premises must be false. The reason why the argument looks compelling is that there are possible, even plausible, situations in which both of the conditional premises are true. And there are even situations in which the third premise is true but the first two premises are assertable (until the truth of the first two premises becomes apparent). We can see the force of this last claim if we invert the order of the premises:

You insult both me and my dog
If you insult me I will not become violent
If you insult my dog I will become violent

∴ Contradiction!

This argument is not compelling at all. Our inclination to accept both the conditional premises when said is undermined by the first premise.

The inclusion of modus ponens in our theory will be useful in chapter 9 below. If we excluded modus ponens, it will make little difference for the theory of conditionals itself. As a fall-back position we can differentiate between conditionals that always allow modus ponens (call them 'ponible conditionals') and those for which there are situations in which modus ponens fails for them. We could then argue that the conditionals I use in chapter 9 to support uses of disjunctive syllogism are ponaible conditionals. But I do not think I need this sort of fall-back position.

7.10 Model theory for conditionals

To reflect the role of the antecedent and consequent in choosing circumstances for the evaluation of a conditional, we use a five-place relation to give a truth condition for the conditional. This relation stands between a pair of propositions (the antecedent and consequent of the conditional) and three situations. Thus, if $C_{|A||B|}stu$, then t and u make up a pair that is reasonably close to s and that satisfy the background conditions that are assigned in s to ' $A \Rightarrow B$ '.

The truth condition itself mimics the Routley–Meyer truth condition for implication, viz.,

' $A \Rightarrow B$ ' is true in s iff $\forall x \forall y ((C_{|A||B|}sxy \ \& \ 'A' \text{ is true in } x) \supset 'B' \text{ is true in } y)$.

We place a few conditions on our relation C . First, in order to satisfy the inference

$$\frac{A \rightarrow B}{\therefore A \Rightarrow B}$$

we need

$$C_{XY}stu \supset Rstu,$$

for all propositions X and Y and all situations s , t , and u . This material conditional might look backwards, but it isn't. It tells us that we only choose circumstances that are already related by R to s .

In order to satisfy modus ponens for the conditional, we use the condition that, for all situations s , and all propositions X and Y ,

$$C_{XY}sss.$$

We can see quite easily that this condition does the trick. For suppose that $A \Rightarrow B$ obtains at s and so does A . Then, by the truth condition for the conditional and the above condition, B is true at s as well.⁴

7.11 The direct argument

Paul Grice has given the following argument to support the identification of the conditional with the material conditional (Grice 1989). Here is a version of Grice's argument. From a disjunction like

Either Prof. Plum or Col. Mustard committed the murder.

⁴ To make modus tollens valid, we need to add the condition that $C_{XY}ss^*s^*$ for all propositions X and Y and all situations s .

we can infer the conditional,

If Prof. Plum did not commit the murder, then Col. Mustard did.

This argument looks valid. Thus, it would seem that inferences of the form

$$\frac{A \vee B}{\therefore \sim A \Rightarrow B}$$

are also valid.

But if we add this inference to our logic of conditionals, our conditional would no longer be relevant. For, suppose that we were to add the ‘direct rule’, viz.,

From $A \vee B_\alpha$ infer $\sim A \Rightarrow B_\alpha$,

to our natural deduction system. We would then be stuck making valid the following argument:

- | | |
|---|------------------------|
| 1. $p_{(1)}$ | <i>hyp.</i> |
| 2. $p \vee q_{(1)}$ | 1, $\vee I$ |
| 3. $\sim p \Rightarrow q_{(1)}$ | 2, direct rule |
| 4. $p \rightarrow (\sim p \Rightarrow q)_\emptyset$ | 1 – 3, $\rightarrow I$ |

So, we would be able to derive a variant on negative paradox.

To retain a relevant theory of conditionals I reject the direct rule. But I follow Stalnaker (Stalnaker 1999) in holding that, whereas the direct rule is not *valid*, it can be *reasonable* to make inferences of this form.

The idea is quite simple. If the participants in a conversation explicitly agree to a statement, such as ‘Either Prof. Plum or Col. Mustard committed the murder’, then it is quite likely that they will constrain the set circumstances they use in that conversation to evaluate conditionals to include only situations that make that statement true. If, in addition, all the circumstances chosen are possible, then the salient conditional will hold.

Let’s be a bit more precise about what is going on here. Suppose that we agree that either Prof. Plum or Col. Mustard committed the murder. And now we want to evaluate the conditional ‘If Prof. Plum did not commit the murder Col. Mustard did.’ We will choose a set of circumstances that is possible (and which satisfies the conditions with regard to both antecedent and consequent), and will include only circumstances in which *both* situations make it true that one or other of the Professor or Colonel is the murderer.

Take one arbitrary such circumstance (t, u) . Suppose that ‘Prof. Plum did not commit the murder’ is true at t . At u , at least one of ‘Prof. Plum committed the murder’ and ‘Col. Mustard committed the murder’ is true (by the truth condition for disjunction). It cannot be the former, since u is co-possible with t and its negation is true at t . Thus, it must be that Col. Mustard is the murderer at u .

Generalising, we have shown that ‘If Prof. Plum did not commit the murder, then Col. Mustard did’ is true.

Note that this argument only holds if, first, the participants in the conversation do in fact constrain the circumstances that they use to include only those that satisfy the disjunction and, second, that it is reasonable for them to choose only possible circumstances to evaluate that conditional. In many (perhaps most) cases in the actual world these conditions are met, and so the inference is a reasonable one.

It is important that we vindicate the direct rule as a reasonable form of inference. We use it often. Suppose, for example, that one is taking a multiple-choice exam. Suppose also that one is asked to name the capital of Kenya and is given the following choices:

- (a) Dar es Salaam
- (b) Nairobi
- (c) Kampala

It would seem that one is correct to say in this situation,

If neither Dar es Salaam nor Kampala is the capital of
Kenya, then Nairobi is. (7.8)

This case is very similar to that of our murder suspects. Here, the disjunctive fact that either (a), (b), or (c) holds is imposed on the situations that we consider when evaluating the conditional. In this context it is inappropriate to look at any other situations – they have been excluded by the rules of the test.

Note that we can even cook up circumstances like this to make true some of the examples that we used to motivate the relevant account. Consider the sentence,

If Jones picks the guinea pig up by the tail, its eyes will
fall out. (7.9)

Suppose that one is taking a test (like an American SAT or GRE) in which one has been asked to read a silly story and then select one of the following answers to test her comprehension:

- (a) Jones does not pick up the guinea pig by the tail.
- (b) The guinea pig’s eyes fall out.

For the same reasons as in our discussion of (7.8), (7.9) is true. This seems right; there is no mitigating difference between the two cases. What is wrong with the guinea pig conditionals is not that there are some contexts of utterance that make them true, but rather that in most contexts they are false. Our theory allows for both. Unlike the classical view, it does not force them always to be true.

7.12 Probability and conditionals

Perhaps the most widely held view about conditionals at present is Adams' thesis. In order to understand Adams' thesis, we first need to understand what it means to say that a sentence is *assertable*. We say that a sentence is assertable if and only if the speaker has sufficient evidence to assert it. Ernest Adams thinks of assertability in terms of subjective probability. A sentence is assertable, on Adams' view, if and only if the speaker attributes to it a high probability. The matter with regard to conditionals, however, is slightly more complicated.

In (Adams 1975), Adams asserts that the probability of a conditional is the same as the corresponding conditional probability. In formal notation, this is

$$\Pr(A \Rightarrow B) = \Pr(B/A).$$

' $\Pr(B/A)$ ' is the probability of B given A . Usually, the conditional probability is determined by the equation $\Pr(B/A) = \frac{\Pr(A \wedge B)}{\Pr(A)}$, when $\Pr(A)$ is not zero (otherwise the conditional probability is undefined).⁵ This equation is often called 'Stalnaker's thesis' because Stalnaker asserted it in (Stalnaker 1970). David Lewis proved, however, that given the usual axioms for probability and the assumption that we can nest conditionals (i.e., put conditionals inside other conditionals), Stalnaker's thesis leads to a sort of triviality. That is, under these conditions Stalnaker's thesis entails that formulae can take only a very small number of values. Thus, Stalnaker's thesis has been abandoned by most philosophers working in this area. With the abandonment of Stalnaker's thesis, many philosophers have adopted the weaker Adams' thesis.

Those who subscribe to Adams' thesis can be placed into two camps. On one hand, there are philosophers like Ernest Adams who hold that there are assertability conditions for conditionals, governed by Adams' thesis, but that indicative conditionals have no truth conditions. On the other hand, there are those like David Lewis and Frank Jackson who believe that conditionals do have truth conditions (in the case of Lewis and Jackson, these are given by the truth table for the material conditional) and also that conditionals have assertability conditions dictated by Adams' thesis.

It is a virtue of Adams' thesis that it gives us reasons to reject contraposition, strengthening of the antecedent, and transitivity. As we have seen, none of these rules are valid for conditionals. And none of them of conditional probabilities. For example, just because $\Pr(C/A)$ is high does not mean that $\Pr(C/A \wedge B)$ is high. Another virtue of Adams' thesis is that it does require some measure of relevance between antecedent and consequent of assertable conditionals. For example, even if it is true that ticket number 1234567 will win the Idaho State Lottery next month, the conditional 'If cats like to eat chicken, ticket

⁵ There are also approaches that take conditional probability to be primitive, which allow $\Pr(B/A)$ to have a value even when $\Pr(A) = 0$.

number 1234567 will win the Idaho State Lottery next month' need not be assertable, since one need not give the corresponding conditional probability a high value.

The notion of relevance captured by the theory of conditional probability, however, is too weak to do justice to our intuitions about conditionals. Dunn's conditional, viz.,

If I scare this pregnant guinea pig, her babies will be born blind.

turns out to be assertable on Adams' thesis, in cases in which the speaker knows that guinea pigs are blind at birth. In this case the probability of the consequent is high and its probability is independent of the probability of the antecedent. This means that regardless of the probability of the antecedent, the probability of the conditional probability, $\Pr(\textit{blind/scare}) = \Pr(\textit{blind})$. Thus, Adams' thesis does not capture our intuitions about relevance. Now, this failure does not by itself imply that Adams' thesis should be rejected. We might add other pragmatic devices to it in order to explain full-blooded relevance. Clearly, however, such devices are needed.

Frank Jackson holds that indicative conditionals have truth conditions given by the truth tables for the material conditional and assertability conditions given by Adams' thesis. A use of 'If ... then', on Jackson's view, carries with it a *conventional implicature*. To understand what a conventional implicature is, consider the word 'but'. The truth table for 'but' is the same as for 'and', but the two words have different connotations. The sentences 'John went to the demonstration and he is tired' and 'John went to the demonstration but he is tired' have rather different tones. The second indicates that there is some apparent tension between the two conjuncts, whereas the first indicates no such thing. According to Jackson, the case of the conditional is very similar. If a speaker states a conditional, then she not only says that the corresponding material conditional is true she also expresses the fact that her corresponding conditional probability is high.

What work, then, is done by the truth condition in Jackson's theory? It has been suggested that the main virtue of the truth condition is that it allows for an analysis of nested conditionals (see e.g. (Woods 1997)). I have doubts about its efficacy in this task and Jackson does too (Jackson 1987). To see why this is so, let us take a closer look.

Supposedly, the problem with the pure Adamsian account is that although it gives us a way of dealing with statements of the form *If A, then B*, where neither 'A' nor 'B' contain conditionals, it cannot treat cases in which they do contain conditionals. For the assertability condition for *If A, then B* is $\Pr(B/A)$'s being high, but we have no assertability condition for, say, *If A, then if B, then C*. Jackson's view, on the other hand, sets the assertability condition for this latter form of statement at $\Pr(B \supset C/A)$'s being high.

So far, so good. But consider that one of the virtues of Adams' thesis is that it captures some of our intuitions regarding the relevance of antecedents to consequents in acceptable conditionals. Now, it is a fact about classical probability theory that $\Pr(\sim B/A)$ is always less than or equal to $\Pr(B \supset C/A)$. So, if the sentence

If petrol fumes are present, I will not strike this match.

is assertable, then so is

If petrol fumes are present, then if I strike this match, I will turn into a cat.

What has happened here is that we have regained all the irrelevance of the material conditional when the conditional we are looking at is within the scope of other conditionals.

These arguments show, I think, that conditional probabilities do not by themselves govern the assertability conditions for conditionals. But it would seem that in many cases a high value for $\Pr(B/A)$ is a good reason for asserting $A \Rightarrow B$. I suggest that this is because, in many cases a high conditional probability is a good indicator of a relevant link between the antecedent and consequent.

7.13 Even if

'Even if' conditionals pose an interesting problem for relevant logicians. 'Even if' is a relevance breaker. Consider the conditional,

I will take Ramsey for a walk even if it rains.

This conditional does not express a very strong link between its raining and my taking my dog for a walk. It says that despite certain tendencies to the contrary, I will walk my dog.

The tendencies to the contrary we will leave aside for a minute. We will concentrate for now on the relevance-breaking property of 'even'. I suggest that when we evaluate an 'even if' conditional, we imagine a set of circumstances. In some of these circumstances, the antecedent is true (if there are any situations in which the antecedent is true). It may be that in all of these circumstances (or in only some or none) that the consequent is true as well. Then, if in every circumstance (t, u) that is considered in which t satisfies 'A', u makes 'B' true, then ' $A \Rightarrow B$ ' is true.

The only difference here between this analysis and our analysis of standard conditionals is that the condition regarding the consequent is dropped. What is happening here is that we are seeing whether in the circumstances that we can

envision (those we think are reasonable) whenever the antecedent is true so is the consequent.

The condition regarding the antecedent is retained. It would seem that vacuity with regard to 'even if' conditionals is as abhorrent as it is with regard to standard conditionals. Consider, for example,

Even if Sally squares the circle you will win your bet.

If you have bet against Sally squaring the circle, then this conditional is false, regardless of whether you and the people to whom you are talking are very reluctant to consider circumstances in which Sally squares the circle.

The other aspect of 'even if's that we mentioned was that they have a negative tone. The assertion of the sentence

Even if it is raining, I will take Ramsey for a walk.

expresses my tendency not to take my dog out when it is raining, or at least to dislike walking a dog in the rain. This negative tone, in my opinion, cannot be dealt with in the semantics for conditionals, but rather is a pragmatic matter; it would seem to be an implicature rather like the negative implicature associated with 'but'.

8 Counterfactuals

8.1 Introduction

Consider again Adams' examples of indicative and counterfactual conditionals:

If Oswald didn't kill Kennedy, then someone else did.

If Oswald hadn't killed Kennedy, then someone else would have.

Someone asserting the first of these sentences would assume that Kennedy was killed, but would not assume that Oswald did it. Someone asserting the second, on the other hand, assumes that Oswald did kill Kennedy.

We have a theory of how to evaluate the indicative conditional. We look at the circumstances that the people involved in the conversation in which it is stated can legitimately envision. If those people are taking this conditional seriously, then they will limit themselves to possible circumstances in which Kennedy is murdered. In those in which Oswald is not the murderer, someone else is. Thus the conditional is true.

I suggest that there isn't a very big difference between counterfactuals and indicative conditionals. I think that the way we evaluate them is very similar. In the case of the Kennedy counterfactual above, what we do is imagine ourselves in a situation in which we have all the salient information but not the information that Kennedy has been murdered (or even the information that he is dead). We then ask whether the indicative conditional 'If Oswald does not kill Kennedy, someone else will' is true in this situation.

The reason that we hesitate about answering this question is that we don't know what all the salient information is. We don't know what the CIA, Cuban Intelligence, the KGB, and the Mafia were planning.

But just because we don't know what the salient facts are, does not mean that there is no fact of the matter about what they are and what circumstances we would consider if we were in a situation determined by those facts. I claim that we can hold that counterfactuals can be true or false regardless of whether we know that they are or even have the means to discover whether they are.

8.2 Sketch of a theory

Let's generalise the process that we described in the previous section. We will use the notation ' $A \Box \Rightarrow B$ ' to mean 'If it were that A , then it would be that B '.

Suppose that we want to evaluate ' $A \Box \Rightarrow B$ ' in the situation s . What we do is look at a set of situations s' that capture the standpoint from which we want to test the relationship between ' A ' and ' B '. In the Kennedy counterfactual examined above, this standpoint is the American and geopolitical landscape of 1963. Consider a very different example:

If Kangaroos didn't have tails, they would fall over.

Here we want to evaluate the relationship between Kangaroos' being tail-less and their falling over in the presence of the actual laws of physics.

Let us call this set that we choose, the set of *base situations*. In the Kennedy example, our base situations will hold fixed the actual facts about the Mafia, Cuba, etc., and in the case of our poor tail-free kangaroos, these situations will include all the laws of physics salient to things falling over or staying upright.

I don't have a complete theory of how we choose our base situations, but there are some points that I can make. First, both the antecedent and consequent, as well as other features of the context of utterance, help to determine the base situations. The role of the antecedent is reasonably clear. We choose base situations that do not contain any information that conflicts with the antecedent of the counterfactual.

The role of consequent is a little more complicated. Here is an example, from Marc Lange (Lange 1993), that we will use several times in this chapter. Suppose that two doctors find an empty syringe beside a patient. The syringe is marked 'cyanide'. One doctor says to the other

If the patient had been injected with that syringe, we would have an excellent journal article.

Here the speaker is choosing base situations in which it is not a law that cyanide is poisonous to people but in which it is assumed that this is a law of nature. And in these base situations, the syringe was filled with cyanide. Then it would be legitimate for the speakers (imagining themselves in these base situations) to choose only circumstances in which that patient is immune to injections of cyanide, but in which people think that cyanide is poisonous to everyone. Using such circumstances to evaluate it, the above counterfactual comes out true.

The second doctor, however, replies:

If the patient had been injected with that syringe, he would be dead.

This too can be counted as true. It is legitimate to choose base situations, like our own, in which, like the doctors' own situation, it is a law that cyanide

is poisonous. In such situations the implication 'The patient is injected with that syringe implies that he dies' holds. This implication in turn implies the corresponding conditional and so the counterfactual is true in the doctors' situation.

Thus, we can see that the consequent of the counterfactual that we are evaluating can help direct us to the base situations for its evaluation. Dov Gabbay is perhaps the first philosopher who noticed the role that the consequent of a counterfactual plays in our choice of situations used to evaluate it. In support of this view, Gabbay asks us to consider the following pairs of counterfactuals:

If I were the Pope, I would have allowed the use of the pill (8.1)
in India.

If I were the Pope, I would have dressed more humbly. (8.2)

and

If New York were in Georgia, then New York would be in (8.3)
the South.

If New York were in Georgia, then New York would be in (8.4)
the North.

An utterance of (8.1) assumes that India retains its problems with overpopulation, whereas an utterance of (8.2) does not seem to assume this at all. In the second pair, an assertion of 8.3 assumes that Georgia remains in the South, but asserting 8.4 assumes that New York remains in the North ((Gabbay 1972), pp. 98f and (Gabbay 1976), p. 188). In our theory, we can cast these assumptions as constraints on the choice of base situations.

There are other factors that constrain our choice of base situations. The purpose for which the counterfactual is uttered can help to determine the correct base situations. Consider a counterfactual used in a scientific explanation. Suppose that we want to explain why a star appears to have moved from its usual location during an eclipse of the sun. The explanation is that because of its location, the mass of the sun attracts the light from the star. This explanation would seem to indicate the truth of the following counterfactual:

If the sun were not near the path of the light from the star to the earth, the apparent position of the star would not have changed.

The base situations chosen to evaluate this counterfactual would include all the relevant actual laws of nature. For it is these laws of nature that are supposed to help explain the apparent shift in the position of the star.

Historical counterfactuals behave similarly. Suppose one wants to argue that Churchill's being prime minister in 1940 saved the United Kingdom from suing for peace with Nazi Germany. He or she might use the following counterfactual:

If Churchill had died in the early 1930s, the UK would have sued for peace with Germany in 1940.

To evaluate this counterfactual, we choose base situations in which every relevant historical factor – Germany’s having overrun Western Europe, the existence of the German–Russian pact, the existence of the appeasement movement in the British government, and so on – except that they would not include Churchill in parliament or even alive (although they would not necessarily have him dead – they just would not include the information that he is alive).

Clearly, in order to develop a theory of our choice of base situations, we would need to do a lot of analysis of how speech contexts work. I hope, however, that the above examples are suggestive enough to indicate to the reader at least a vague picture of how such a theory would look.

8.3 Counterpossible conditionals

In chapter 1 above, a motivation that I gave for relevant logic was the ability to use it (and its semantics) to develop a theory of counterpossible conditionals. Consider again the statement

If Sally were to square the circle, we would be surprised.

Squaring the circle (with the use of only compass and straight-edge) is impossible. Thus this is a counterpossible conditional.

Counterpossible conditionals provide no special problem for us. In the evaluation of conditionals we have chosen base situations which do not contain any information that conflicts with the antecedent of the counterfactual. To evaluate the counterfactual above, it would seem that we should do the same. We can choose base situations which do not contain the information that it is impossible to square the circle. But we choose situations in which we think it is impossible to do this task. We then choose circumstances that are appropriately connected to our base situations and evaluate the corresponding conditional, ‘If Sally does square the circle, we would be surprised.’¹

8.4 Counterlegal conditionals

Treating counterlegal conditionals is also reasonably straightforward. Consider a counterfactual with an antecedent that says that an actual law of nature is violated. The following is the sort of counterfactual that one sees in ‘fine-tuning’

¹ Note that the indicative conditional ‘If Sally does square the circle, we will be surprised’ also comes out true on a normal choice of circumstances to evaluate it. I leave it as an exercise for the reader to convince him or herself of this.

versions of the cosmological argument for the existence of God:

If the gravitational constant were even slightly different, the existence of living things would be impossible.

In evaluating this counterfactual, we choose base situations in which other laws of nature are the same as the actual ones, but which do not contain the information that the gravitational constant is what it actually is. This becomes difficult if the other laws of nature somehow entail that the gravitational constant has a particular value. If this is the case, then we look at situations in which the laws are slightly different from our own so as not to contradict the antecedent.

Not all counterlegals are counter-legal because of their antecedents. Let's return again to Lange's example of the two doctors and the syringe. The first counterfactual – 'If the patient had been injected with this syringe, we would have an excellent journal article' – is a counterlegal. Why should we count it as a counterlegal? It is clear from the statement that it means to ask us to think about what would happen if the *apparent* laws of nature are false. In this case, the apparent laws of nature are also the real laws of nature. So, it is a counterlegal.

If the antecedent and consequent of a counterfactual are not jointly nomically possible, then we should suspect that we are dealing with a counterlegal. But this isn't always the case. Consider the counterfactual,

If I drop this pen, the gravitational constant will change.

It would seem that this statement is not only false, but it would not seem even to be a counterlegal. The difference from the previous example is that in the syringe case the speaker is focusing on another fact about the present situation, the fact that the patient is still alive. If we fix that fact in all the circumstances used to evaluate the conditional and include some circumstances in which the patient is injected with the syringe, then we will be looking at circumstances that do not obey our laws of nature. In order to use such circumstances, we need to choose base situations that do not include all our laws of nature. In the case of the pen-dropping counterfactual, there would seem to be no feature of the situation, which the consequent indicates, that could be used in the same way.

8.5 Laws of nature and counterfactuals

We have now touched on the relationship between counterfactuals and laws of nature. This relationship is multi-faceted. We will discuss this relationship in this and the following two sections.

In several papers published in the 1950s, Nelson Goodman and Roderick Chisholm argued that one feature that distinguishes statements of laws of nature from other statements is the relationship that the former bear to counterfactuals.

Statements of laws of nature, or those statements purporting to express laws of nature, somehow indicate the truth of certain counterfactuals. Thus, to use another old chestnut, the general statement

There are no balls of pure uranium 235 that are one mile in diameter. (8.5)

indicates the truth of the counterfactual

If there were a ball one mile in diameter, it would not be made of uranium 235.

On the other hand, the general statement

There are no balls of pure gold one mile in diameter. (8.6)

does not indicate the truth of the counterfactual

If there were a ball one mile in diameter, it would not be made of pure gold.

The fact that (8.5) indicates the truth of the corresponding counterfactual (8.6) does not, according to Goodman and Chisholm, is explained by the fact that the former expresses a law of nature whereas the latter does not.

Our theory of counterfactuals explains the Chisholm–Goodman property of laws rather well. When we look at the relationship between a law of nature and a counterfactual, we choose base situations that make the law true. Since laws are informational links, they constrain which implications are true in these situations and so the relevant counterfactuals come out true. When we evaluate the counterfactual, ‘If there were a ball one mile in diameter, it would not be made of uranium 235’, we look at base situations that agree with the actual situation about the relevant laws of nature. Thus, if we go to one of the base situations we have chosen, we will have available to us only circumstances that are such that if an attempt is made to assemble a ball of U235 that is one mile in diameter, it will explode. If we choose only nomically possible circumstances, as we should according to our theory, then in any such in which there is a ball one mile in diameter, it is not made from U235. Thus, the counterfactual comes out true.

When we look at the accidental generalisations, we cannot give the same argument. We would need a reason, even if we choose base situations in which the generalisations are true, to examine only circumstances in which the generalisation holds. Only in special contexts would there be such reasons. On the other hand, it is rather standard to look at base situations in which our laws hold. In fact, we need reasons not to do so when evaluating counterfactuals.

Now, some philosophers have suggested that there is a much closer relationship between laws of nature and counterfactuals. Among them are John Pollock and John Carroll. Pollock proposes a principle of *legal conservativeness*

((Pollock 1984), p. 118 and see (Carroll 1994)). Pollock assumes a comparative similarity semantics for counterfactuals. His principle of legal conservativeness says that in evaluating a counterfactual at the actual world, we find, if possible, worlds in which the actual laws of nature hold. Where the antecedent of the counterfactual conflicts with the laws of nature, we find worlds in which the laws of nature are minimally changed to be consistent with the antecedent. As we have seen, I disagree with Pollock. I think that it is sometimes legitimate to hold fixed other features of the context and use base situations in which our laws of nature fail to hold.

There are good and bad arguments in the literature against legal conservativeness. As I have said, I think that Lange's example provides us with a good argument against it. An unsound, but interesting, argument is due to John Halpin (Halpin 1991). The following counterfactual seems plausible enough:

If Caesar had lead the United Nations in Korea, he would have used the same cunning that he used in Gaul.

If, however, Caesar had been alive during both the Korean War and during the Roman invasion of Gaul, he would be well over two thousand years old. And this seems to be a violation of the laws that cause our bodies to degenerate. Thus, Halpin claims that these laws of nature are not among the background conditions for this counterfactual.

I am not convinced by this argument. What the counterfactual is saying is that Caesar, if he were in charge in Korea, would have used the same cunning that he *actually* had used in the Roman invasion of Gaul. Compare the following counterfactual:

If Jill hadn't started smoking when she was young, she would be much taller now than she is.

Now, this counterfactual does not require for its truth a situation in which Jill has two distinct heights at the same time. Instead, it makes a comparison across situations, between an actual situation and a non-actual situation.

A useful theory that cashes out this intuition is to be found in chapter 16 of (Cresswell 1990). There Cresswell claims that comparative statements commit us to *degrees*. For example, the statement

I am a better philosopher than a bridge player.

is true if and only if the degree of my goodness as a philosopher is greater than the degree of my proficiency at bridge. In light of Cresswell's theory, we can now understand Halpin's counterfactual. As I have said, I think that a fair paraphrase of it is:

If Caesar had fought in Korea, he would have done so with the same cunning as he actually did in Gaul.

This counterfactual is true if and only if there is a degree of cunning which Caesar had applied in his invasion of Gaul, and this degree of cunning is matched by Caesar in the relevant hypothetical situations in which he fights in Korea.

Although I reject legal conservativeness as a universal rule, I agree with the intuition behind it to a certain extent. Normally, we look at base situations in which our laws hold as laws. But we have to be sensitive to context and the content of a counterfactual statement to determine whether or not to choose base situations that have our laws.

Bas van Fraassen, on the other hand, has argued that there is a much looser relationship between laws and counterfactuals than this. His argument is given in (van Fraassen 1981), (van Fraassen 1980), and (van Fraassen 1989). The following is a concise presentation of this argument:

The truth-value of a [counterfactual] conditional depends in part upon the context. Science does not imply that the context is one way or another. Therefore science does imply the truth of any counterfactual – except in the limiting case of a conditional with the same truth value in all contexts. ((van Fraassen 1980), p. 118)

The argument is quite straightforward. A law of nature is true in every context. But a counterfactual may be true in some contexts and not in others. Thus, laws of nature do not imply counterfactuals, except in a rather uninteresting sense. We should note a few things about this argument in relation to my own view. First, it is not my view that laws of nature are true in every context, where ‘context’ is taken to mean ‘situation’. Second, ‘implication’ here, I think, should be taken to mean ‘logical implication’ or ‘entailment’.

Let’s leave aside the issue of whether laws of nature are true in all contexts and assume with van Fraassen that, if there are any laws of nature they are true in all contexts. Clearly, van Fraassen’s argument is valid. But what it shows is that laws of nature do not entail ordinary counterfactuals. I don’t think that they do either. But van Fraassen thinks that we can draw a much deeper conclusion than this:

There was at one point a hope, expressed by Goodman, Reichenbach, Hempel, and others, that counterfactual conditionals provide an objective criterion for what is a law of nature, or at least, a lawlike statement (where a law is a true lawlike statement). A merely general truth was to be distinguished from a law because the latter, and not the former, implies counterfactuals. This idea must be inverted: if laws imply counterfactuals then, because counterfactuals are context-dependent, the concept of law does not point to any objective distinction in nature. (Ibid.)

So, van Fraassen thinks that this argument indicates a flaw in the concept of law, that this concept does not pick out anything 'objective in nature'. If so, this is rather depressing news for supporters of laws of nature, like myself.

The problem with van Fraassen's argument is that he assumes that the relationship between laws and counterfactuals, if there is one, must be that of entailment. I don't hold this view, nor, I think, do most other philosophers who have written about counterfactuals. Goodman thought that laws were 'connecting principles' between the antecedent and consequent of a true counterfactual, but he did not claim that they entail counterfactuals. In fact, he indicated that other truths are usually needed to support counterfactuals. Likewise, I hold that in standard cases we look for base situations that agree with our own situation on the laws of nature, but I do not claim that these laws are *alone* responsible for making counterfactuals true.

8.6 Lewis and miracles

Before we leave the topic of laws and counterfactuals, we should discuss perhaps the most famous argument against legal conservativeness. This is David Lewis' argument from (Lewis 1979). The difficulty for Lewis centres around the problem of backtracking counterfactuals. Consider the counterfactual,

If I slept through my alarm this morning,
I would have stayed up late last night.

Here we have a backtracking counterfactual. The antecedent is made true by events (in hypothetical situations) that occur after those that make the consequent obtain. According to Lewis, we standardly avoid backtracking in counterfactuals; we allow antecedents to make a difference to what will happen, but not to what has happened. That is, in evaluating counterfactuals, we usually take the past (before the time indicated by the antecedent) as being the same as the actual past.

Lewis' concern in (Lewis 1979) is to add principles to his similarity semantics for counterfactuals such as to make our standard asymmetrical attitude to past and future come out right in most cases. Lewis considers counterfactuals starting with the antecedent 'If Nixon had pushed the button'. Suppose for a moment that the actual world is governed by deterministic laws. Then, if we keep the laws of nature fixed, Nixon could only have pressed the button if the past had been different, perhaps very different.

Lewis' view, however, is that when evaluating counterfactuals with this antecedent, we should not look at worlds in which the past was different. He thinks that we should apply our standard tendency to allow more variation in the future than in the past in choosing the indices that we use to evaluate counterfactuals.

He does not think that the following counterfactual is true:

If Nixon had pressed the button, the past would have been very different.

To reject this counterfactual, Lewis also rejects legal conservativeness. The worlds which we look at to evaluate the counterfactual do not obey the same laws of nature as our world. We look at worlds in which the past up to the point at which Nixon is suppose to have pushed the button is the same as our world, but then there is an event that violates the laws of our world but which causes Nixon to push the button (or perhaps is in fact just his pushing the button). From the point of view of our world, this event is a miracle. Lewis allows miracles of this sort, to allow fit between our world and worlds used to evaluate counterfactuals.

Jonathan Bennett thinks that Lewis' worry about backtracking is unwarranted. Consider the following example (Bennett 1984):

If Stevenson had been President in 1953, he would have won the election in 1952.

This statement is true on Lewis' theory, but it is a backtracking counterfactual. A big miracle would have been necessary to make Stevenson president if he hadn't won the election in 1952. A much smaller miracle, plus a bit of backtracking, is preferable according to Lewis' theory. And this preference seems satisfied by the counterfactual above.

Bennett claims, moreover, that a lot of the counterfactuals that we count as true require backtracking. 'If there had been twice as many Jews in Germany in 1933, then there would have been a baby boom in the Jewish community in the nineteenth century' seems true, but it is again a backtracking counterfactual. Bennett points to the fact that we often accept backtracking counterfactuals to undermine Lewis' view that such cases are non-standard. Bennett says that Lewis' theory cannot stop the 'floods of backward counterfactuals' and so '[t]here seems to be nothing left of the supposed (counterfactual) foundation for time's arrow' (p. 79).

The importance of Bennett's argument is that it purports to show that Lewis' worries do not really play against legal conservativeness. Lewis rejects legal conservativeness because it forces us to look at worlds in which the past is very different from that of the actual world in a great many cases, thus violating the supposed standard approach to counterfactuals. Bennett claims that Lewis' theory itself forces us to look at 'backtracking worlds' in a great many cases, and so Lewis' theory does not save the intuitions that motivated its creation. Bennett's response to all of this is to abandon the standard approach and embrace legal conservativeness.

As I have said, I reject legal conservativeness in its strong form. In some cases, we can legitimately use base situations in which not all of our laws of nature hold even when the antecedent of the counterfactual is not explicitly counterlegal. But these cases are relatively rare. Most of the time we do use base situations that contain all our laws of nature.

What allows us to avoid worries like Lewis' is that our situations, as opposed to his worlds, are very partial. When we evaluate the Nixon counterfactual we use base situations in which all the *salient* information is the same. Nixon is president, the button is connected to the nuclear arsenal in the way that it was in the early 1970s, the Soviet missile launch system is also the same, and so on. When we choose circumstances to evaluate the counterfactual, we do not look for ones that have an entire history as like the base situations as possible. They might be very limited in terms of the information that they contain. Since we are not looking for the best overall fit between entire histories, the supposed trade-off between violations of laws of nature and this overall-fit does not arise.

We can agree with Bennett to a large extent. When we are using counterfactuals such as 'If there had been twice as many Jews in Germany in 1933, then there would have been a baby boom in the Jewish community in the nineteenth century', we are intending to make claims about the past. So we look at circumstances that include information about the past. They don't need to include information about the whole history of the universe, for we usually only make claims about very limited stretches of history.

8.7 Countermathematicals

In the introductory chapter above, I presented a counterfactual due to Hartry Field, that is,

If the axiom of choice were false, the well-ordering lemma would fail.

The well-ordering lemma, in Zermelo–Frankel set theory (ZF), is equivalent to the axiom of choice. Let us suppose that this is a countermathematical conditional; let us assume that the axiom of choice is true. Field thinks that there is no such thing as metaphysical necessity, but I do. I will assume that all mathematical truths are also metaphysically necessary. If this assumption is false, then it will be even easier to have my view accommodate this counterfactual.

When we evaluate this conditional, it is natural to look at base situations in which the other axioms of ZF hold, but not the axiom of choice. Let us consider one such base situation, *s*. From *s* we choose a set of circumstances. Each of these will satisfy the other axioms of ZF. Some will make the axiom of choice

false. It is natural also to choose circumstances in which the following formula is true:

$$AC \equiv WO,$$

where 'AC' is an abbreviation of the axiom of choice, 'WO' is an abbreviation of the well-ordering theorem, and ' $\equiv$ ' is material equivalence.²

So far we are very nearly repeating the procedure that we used when treating the direct argument. But there is a slight difference. We can't choose only pairs of situations that are jointly possible relative to our base situations. Some of our base situations will be possible, but they will have to 'look at' situations in which the axiom of choice is false, and these latter situations are impossible. What we do is look at pairs of situations that contain no contradictions in the language of set theory, nor do they contradict each other in this vocabulary.

Now we can see that Field's counterfactual comes out true. Let (t, u) be a circumstance used to evaluate it. As we have seen, ' $AC \equiv WO$ ' is true in each of t and u . Let us suppose that $\sim AC$ obtains in t . Since t contains no contradictions, $\sim WO$ is also true in t . Since ' $AC \equiv WO$ ' is true in u , ' $AC \vee \sim WO$ ' obtains in u too. No sentence in the language of set theory at u contradicts any at t . So AC does not hold in u . Thus, by the truth condition for disjunction, $\sim WO$ obtains in u . Generalising, the counterfactual,

$$\sim AC \Box \Rightarrow \sim WO$$

is true, as we set out to show.

We can use the same technique to deal with Barwise's paradox. In chapter 5 of (Barwise 1989), Jon Barwise tells the following story. Suppose that there are two mathematicians, Virgil and Paul. Virgil has been trying to prove Goldbach's Conjecture (GC). Virgil also believes in a standard treatment of counterfactuals that uses possible worlds. Paul has been working on another conjecture, called 'RH'. Virgil then discovers that GC and RH are equivalent. Virgil makes the following two claims to Paul:

If you could give a correct proof of RH, I could give a correct proof of GC. (8.7)

It is false that if you could give a correct proof of RH, I could give a correct proof of not-GC. (8.8)

These both seem right. But Virgil then concludes from (8.7) and (8.8) that GC is true. He reasons as follows. From (8.8) there must be some possible world

² Material equivalence is defined as follows: $A \equiv B \text{ =}_{df} (A \supset B) \wedge (B \supset A)$.

in which RH is true and not-GC is not true. Thus, in this world GC is true. But if GC is true in any possible world, it is true in all possible worlds. Thus, Goldbach's Conjecture is true. In this way, we can show that any mathematical statement is true. All we need is some statement that is materially equivalent to it and these are not hard to find.

When we evaluate (8.7), we choose a set of base situations in which $GC \equiv RH$ obtains. When we locate ourselves in one of these base situations, we look at a set of circumstances in which this equivalence also holds. By the same reasoning as we used to evaluate the Field countermathematical, we can derive that statement (8.7) is true. Now, let us turn to the evaluation of (8.8). If we also use a set of base situations in which this equivalence is true and from their use sets of circumstances in which it is true, we find that the counterfactual

If you can give a correct proof of RH, then I could give a correct proof of not-GC.

is not true. For we choose sets of circumstances some of which make RH true. And we would look at sets of pairs of situations that do not contradict themselves or each other in the language of mathematics. Thus, since the equivalence is also true, not-GC is false (and not also true).

Now, we have shown that the above counterfactual is not true. Is its negation true? If we are in a worldly situation, then it is, since worldly situations are bivalent. Otherwise, we should be content with showing that it is not true. I suggest that it does little harm to Barwise's intuitions to claim that the negation in statement (8.8) is really an expression of a denial rather than a real negation.

8.8 Consequent conjunctivitis

One feature of our theory of counterfactuals is that it makes the following inference invalid:

From $A \Box \Rightarrow B$ and $A \Box \Rightarrow C$ to $A \Box \Rightarrow (B \wedge C)$

We will call this rule 'conjunction in the consequent'.

The rule of conjunction in the consequent might seem reasonable at first glance. For suppose that the leader of the opposition in parliament tells you that he would have spent more money on health care and education if he were elected and that he would have also lowered taxes if he were elected. You have a right to infer that he is telling you that he would have both spent more money on health care and education as well as lowered taxes if he were elected. It seems that in certain cases at least that we take for granted that conjunction in the consequent governs our logical commitments.

But this rule is not generally acceptable. For consider again our doctors discussing the syringe and their two statements:

If the syringe had contained arsenic, the patient would have died.

If the syringe had contained arsenic, we would have had a good publication.

Clearly, we do not want to infer from these two statements that

If the syringe had contained arsenic, the patient would have died and we would have had a good publication.

Thus, at least for counterfactuals, conjunction in the consequent does not look like a valid rule.

Interestingly, depending on how we develop the theory of indicative conditionals the rule of conjunction in the consequent might in fact hold for them. But it might not seem so at first glance since, like the counterfactual, the consequent plays a large role in determining how a conditional is evaluated.³ But consider two conditionals ' $A \Rightarrow B$ ' and ' $A \Rightarrow C$ ', both true in a situation s . We might describe the process of evaluating them as follows. First we consider the class of circumstances that we are immediately disposed to consider. Then (if necessary) we enlarge this set to include circumstances (whose first member) makes true ' A '. Next, (if needed) we enlarge the set to include some circumstances in which ' B ' does not hold (in their second members). By our assumption that ' $A \Rightarrow B$ ' is true in s , then for all circumstances in this set in which ' A ' is true in the first member, ' B ' is true in the second member.

Now we turn to evaluating ' $A \Rightarrow C$ '. We do the same thing again, starting with our default set of the circumstances that we are immediately prepared to consider. Now, let us assume that when we look at larger and larger sets of circumstances our ordering of these sets is not affected by the consequent of the conditional. In the terminology of formal semantics, our ordering of sets of circumstances is *nested*. Then we will have a case in which the circumstances we are considering with regard to ' $A \Rightarrow B$ ' is either a superset or a subset of those that we are considering with regard to ' $A \Rightarrow C$ '.

Now let us turn to examine ' $A \Rightarrow (B \wedge C)$ '. Our process of evaluation here is the same. We start with the same default set of circumstances and enlarge it until we have satisfied the constraints concerning the antecedent and consequent. If the sets we consider are nested, then the class of circumstances that we consider for ' $A \Rightarrow (B \wedge C)$ ' will be a subset of both the sets that we consider for ' $A \Rightarrow B$ '

³ Actually, until I had tried to provide a counterexample to conjunction in the consequent for the indicative conditional, I hadn't realised that it was valid.

and ' $A \Rightarrow C$ '. By the truth condition for the conditional and the truth of these two conditionals, it follows that ' $A \Rightarrow (B \wedge C)$ ' also holds in s .⁴

8.9 Modus ponens again

Usually we state counterfactuals only when we think the antecedent is false. But what happens when we find out that the antecedent is actually true? Suppose that one of my colleagues is, unknown to me, slumming. Actually, she is a millionaire. I say, 'If Ismay were a millionaire she wouldn't be working with us.' After I have made this claim, Ismay reveals her secret and tells us that she is in fact rich. Should I reject my counterfactual, or merely say that its truth is independent of the truth of its antecedent?

I would say that we should reject the counterfactual. Thus, it would seem that we have the inference from

$$A \wedge \sim B$$

to

$$\sim(A \Box \Rightarrow B).$$

We can get this inference by accepting modus ponens in the following form:

$$\text{From } A \Box \Rightarrow B_\alpha \text{ and } A_\alpha \text{ to infer } B_\alpha.$$

We can also get it by accepting modus tollens in a very similar form:

$$\text{From } A \Box \Rightarrow B_\alpha \text{ and } \sim B_\alpha \text{ to infer } \sim A_\alpha.$$

I suggest that we accept both of these inferences, for they both seem intuitively reasonable.

⁴ More formally, we have

$$C_{|A||B \wedge C|stu} \supset C_{|A||B|stu}$$

and

$$C_{|A||B \wedge C|stu} \supset C_{|A||C|stu}$$

for all t and u . Thus, from

$$\forall x \forall y ((C_{|A||B|sxy} \& x \models A) \supset y \models B)$$

and

$$\forall x \forall y ((C_{|A||C|sxy} \& x \models A) \supset y \models C)$$

we can infer

$$\forall x \forall y ((C_{|A||B \wedge C|sxy} \& x \models A) \supset y \models B \wedge C).$$

Thus $s \models A \Rightarrow (B \wedge C)$.

There is a simple semantic condition that makes valid modus ponens. That is, if the antecedent is true at s , then s is itself a base situation used to evaluate the counterfactual. Making modus tollens valid is just as easy. We merely constrain our models to say that if the negation of the consequent holds at s , then s is a base situation.

8.10 Counterlogicals

The problem of counterlogical conditionals requires a very different approach. A counterlogical is a conditional the antecedent of which is equivalent to the negation of a law of logic or states that some rule of logic fails. Consider the following counterpossible antecedents:

If modus ponens were to fail, . . .

If intuitionist logic were right, . . .

To make sense of conditionals with these antecedents it would seem that we need situations that are not closed under the principles of relevant logic. For example, to evaluate statements beginning with ‘If intuitionist logic were right’, it would seem that we require situations not closed under the intuitionistically invalid form of double negation.

Now, given our ontology, it is quite easy to construct models for intuitionist logic or models for logics in which modus ponens fails at some situations.⁵ But the ‘pseudo-situations’ in these models do not belong to our intended model. In sense, then, these pseudo-situations are not used to interpret the logical operators of natural language. Instead, they are used to interpret some of our metalinguistic assertions about non-natural languages.⁶

Here is the principle that we use to demarcate statements that are to be interpreted in terms of situations closed under the principles of relevant logic and those for which we may use ‘non-relevant’ situations:

We may use non-relevant pseudo-situations to interpret a (counterfactual or indicative) conditional only if the antecedent of that conditional is properly to be read as metalinguistic.

One might argue, however, that it is not always clear – there may even be no fact of the matter in certain cases – when an antecedent is metalinguistic. This might be true, but there are clear cases on either side of the divide. ‘If intuitionist logic

⁵ As long as the logic has models of this form. For logics with just, say, algebraic models, it is less clear what it means to suppose that they are correct, without some further qualification on the notion of correctness in use.

⁶ Not that it is inconceivable that some society use intuitionist logic, say, as the inference engine for their natural language. The hypothesis here is that none actually do.

were right' is uncontroversially metalinguistic. On the other hand, 'If Sally were to square the circle' is clearly not metalinguistic. That there are clear cases on either side seems to speak for the need for this demarcation and, given that there are some non-metalinguistic counterpossibles, our original motivation for relevant logic stands.

8.11 Summary of part II

In the last two chapters we have developed a theory of indicative conditionals and a theory of counterfactual conditionals. I have argued that, whereas neither the indicative nor counterfactual conditional is identical to relevant implication, both should be treated as being similar to relevant implication. A semantics based on the model theory for relevant logic is given for the indicative conditional and made slightly more complicated to treat counterfactuals.

Part III

Inference and its applications

9 The structure of deduction

9.1 Introduction

This chapter's purpose is primarily to give us some tools to use in the next two chapters. It will define the technical term 'theory' and it will discuss two sorts of logical consequence. But it will not only introduce these concepts. Each of them is philosophically interesting and some are connected with debates in the philosophical literature. So, we will stop to consider each idea carefully.

Throughout this book I have taken relevant logic to be primarily about inference. This might seem like an odd thing to say, but not every philosopher of logic takes inference to be primary. Some seem more concerned with what theorems the logic has.

This difference in concern might not seem that important. For, the theorems of a logic are connected to the valid inferences by the deduction theorem. We discussed the deduction theorem in chapters 2 and 3 above. There we discussed its relation to our semantics. Here we will discuss it in more proof-theoretic terms.

To make things easier, we will write deductions horizontally. That is, we will write

$$A_1, \dots, A_n \vdash B$$

instead of

$$\frac{\begin{array}{c} A_1 \\ \vdots \\ A_n \end{array}}{\therefore B}$$

The symbol ' $\vdash$ ' is called the 'turnstile'. A structure of the form ' $A_1, \dots, A_n \vdash B$ ' is called a *sequent*. In the 1930s, Gerhard Gentzen developed various logical systems of sequents, known as *sequent calculi*. We will not deal directly with sequent calculi in this book, although various such calculi have been developed for relevant logics by J. M. Dunn, Gregor Mints and Nuel Belnap (see (Anderson and Belnap 1975) and (Anderson *et al.* 1992)). Instead, we will merely discuss what it means for a sequent to be valid in relevant logic.

The sequents we discuss here can have zero or more premises (although only a finite number of them). A sequent with no premises, such as $\vdash A$, tells us that something is a theorem of our logic, R (in this case A). Every one of our sequents will have exactly one conclusion. It might seem odd to mention this, but in some of Gentzen's calculi, more than one conclusion is allowed, and in every one of these calculi there are sequents with no conclusions. We will deal briefly with sequents with more than one conclusion later, but for now we can ignore them.

The deduction theorem is a key to understanding what sequents mean. In its most general form, the general version of the deduction theorem is:

Deduction theorem for a logic L . If $A_1, \dots, A_n, B \vdash C$ is a valid sequent in L , then so is $A_1, \dots, A_n \vdash B \rightarrow C$.

This means that if we can derive C from the premises $A_1, \dots, A_n$ and B , then we can derive $B \rightarrow C$ from $A_1, \dots, A_n$. The converse of the deduction system is also valid. It says

Converse deduction theorem for a logic L . If $A_1, \dots, A_n \vdash B \rightarrow C$ is a valid deduction in L , then so is $A_1, \dots, A_n, B \vdash C$.

We will see soon that the deduction theorem has a very close connection with the rule of implication introduction and the converse deduction theorem has a similar connection with the rule of implication elimination.

In order to understand what the deduction and converse deduction theorems tell us about inferences in relevant logic, we will attempt to answer the following questions.

- What is the relationship between sequents and natural deduction proofs?
- What does the comma between premises mean?

9.2 Sequents and natural deduction inferences

The relationship between sequents and natural deduction inferences is straightforward, at least on the surface. The sequent,

$$A_1, \dots, A_n \vdash B,$$

is valid if and only if we can derive B relevantly from the hypotheses $A_1, \dots, A_n$. In other words, this sequent is valid if and only if given the assumption of $A_1, \dots, A_n$ we can derive $B_{[1, \dots, n]}$.¹

¹ In this chapter, we will relax our conventions regarding quotation marks. Putting in every required quotation mark would make this chapter very difficult to read.

This correlation between sequents and natural deduction inferences helps us to understand further the deduction and converse deduction theorems. For suppose that we have as a valid sequent, $A_1, \dots, A_n, B \vdash C$. Then we know from the correlation that the inference from $A_{1\{1\}}, \dots, A_{n\{n\}}$ and $B_{\{n+1\}}$ to $C_{\{1, \dots, n+1\}}$ is valid. Thus, by implication introduction, we derive that it is valid to infer $B \rightarrow C_{\{1, \dots, n\}}$ from $A_{1\{1\}}, \dots, A_{n\{n\}}$. Running our correlation the other way, we derive that $A_1, \dots, A_n \vdash B \rightarrow C$ is valid.

The case of the converse deduction theorem is just as easy. Suppose that we are given that $A_1, \dots, A_n \vdash B \rightarrow C$ is valid. Then we know, by the correlation, that the natural deduction inference from $A_{1\{1\}}, \dots, A_{n\{n\}}$ to $B \rightarrow C_{\{1, \dots, n\}}$ is also valid. We then ask ourselves what we could derive if we assume $B_{\{n+1\}}$ as a hypothesis along with $A_{1\{1\}}, \dots, A_{n\{n\}}$. By implication elimination, we know that from $B \rightarrow C_{\{1, \dots, n\}}$ and $B_{\{n+1\}}$ we can infer $C_{\{1, \dots, n\} \cup \{n+1\}}$ or $C_{\{1, \dots, n+1\}}$. Thus, we derive that the inference from $A_{1\{1\}}, \dots, A_{n\{n\}}$ and $B_{\{n+1\}}$ to $C_{\{1, \dots, n+1\}}$ is valid. By our correlation then, we know that the sequent $A_1, \dots, A_n, B \vdash C$ is also valid.

The relationship between the sequents and the natural deduction system is not entirely as straightforward as this, especially when we consider other connectives and their relationships to one another. But we will avoid these topics for the time being and keep everything reasonably straightforward.

9.3 What is a premise?

Let us turn now to the second question.

In classical and intuitionist logic, we can take the collection of premises in a sequent to constitute a *set*. But not in relevant logic.² We will show why they cannot be considered to make up a set by assuming that they do constitute a set and seeing what goes wrong. We start by considering a very trivial sequent:

1. $A \vdash A$.

Since the premises are supposed to be a set, we can rewrite this sequent as,

2. $\{A\} \vdash A$.

But, in sets it does not matter how many times an item occurs. In particular, we have the following equation:

3. $\{A, A\} = \{A\}$.

So, from 2 and 3, we obtain $\{A, A\} \vdash A$ or,

4. $A, A \vdash A$.

² But see section 9.7 below.

Now we can use our deduction theorem to get,

$$5. A \vdash A \rightarrow A$$

and again to derive,

$$6. \vdash A \rightarrow (A \rightarrow A).$$

Now, $A \rightarrow (A \rightarrow A)$ might not look like something that a relevance logician should despise, but adding it to the logic R has nasty consequences. It allows us to derive paradoxes of implication such as,

$$(A \rightarrow B) \vee (B \rightarrow A).$$

Adding this paradox substantially weakens the notion of implication. If for any two propositions, one implies the other, implication is not a very strong relation. And it is certainly not the relation that we motivated in chapter 3 above.

Thus, we abandon the notion that our premises make up a set. Rather, they make up what is called a 'multiset'. Multisets have certain properties in common with sets. They are unordered – their elements do not come in a sequence. They are extensional – multisets with exactly the same members and the same number of occurrences of the same members are the same multiset. But the difference between sets and multisets is that two multisets can differ just in the number of occurrences of a given member. Thus (A, A, B) is a different multiset from (A, B) and from (A, A, A, B) but not from (A, B, A) or (B, A, A) .

We have some insight now into the meaning of the comma between premises. The comma joins two premises in the same way that commas distinguish between the members of a multiset. To make this understanding more 'logical', we look at the relationship between the comma and conjunction.

9.4 Intensional conjunction

In classical logic, we have the following equivalence. $A_1, \dots, A_n \vdash B$ is a valid sequent if and only if $A_1 \wedge \dots \wedge A_n \vdash B$ is a valid sequent. This seems rather intuitive. But we cannot accept it in relevant logic. To see why, we note that it is valid in relevant logic that,

$$1. A \wedge B \vdash A.$$

If we have the equivalence between the comma and conjunction, we then can derive,

$$2. A, B \vdash A.$$

By the deduction theorem we get,

$$3. A \vdash B \rightarrow A,$$

and using the deduction theorem again, we obtain,

$$4. \vdash A \rightarrow (B \rightarrow A).$$

This, of course, is positive paradox. Thus, we must reject the idea that when we state a multiset of premises, we are actually stating their conjunction.

In order to represent valid deductions within the logical language, relevant logicians have introduced another connective, which is called *intensional conjunction* or *fusion*. The fusion of two formulae A and B is written $A \circ B$ (read ‘ A fuse B ’). We then replace the classical equivalence between the comma and conjunction with an equivalence between the comma and fusion. That is, in relevant logic, $A_1, \dots, A_n \vdash B$ is a valid sequent if and only if $(A_1 \circ \dots \circ A_n) \vdash B$ is a valid sequent. In order to block the above argument going through for intensional conjunction, in relevant logic the sequent $A \circ B \vdash A$ is not valid.³

From the equivalence of $A_1, \dots, A_n \vdash B$ and $(A_1 \circ \dots \circ A_n) \vdash B$, by multiple applications of the deduction theorem and converse deduction theorem, we can see that all of the following are also equivalent:

- $A_1, \dots, A_n \vdash B$;
- $(A_1 \circ \dots \circ A_n) \vdash B$;
- $\vdash (A_1 \circ \dots \circ A_n) \rightarrow B$;
- $\vdash A_1 \rightarrow (\dots (A_n \rightarrow B) \dots)$.

We can see, then, that there is a very close connection between fusion and implication. In our logic R, we can define a connective that has this property in terms of negation and implication, viz.,

$$A \circ B =_{df} \sim(A \rightarrow \sim B).$$

If we substitute material implication for relevant implication, this is a definition of extensional conjunction ($\wedge$). Thus, by analogy, we can justify interpreting fusion as an intensional form of conjunction.

The connection between fusion and implication is even stronger. The following is a valid formula of R:

$$((A \circ B) \rightarrow C) \leftrightarrow (A \rightarrow (B \rightarrow C)).$$

In semantic terms, this means that for any situation s , ‘ $(A \circ B) \rightarrow C$ ’ is true at s if and only if ‘ $A \rightarrow (B \rightarrow C)$ ’ is also true at s . Given the theory of situated implication of chapter 3 above, we can read ‘ $(A \circ B) \rightarrow C$ ’ as meaning ‘if ‘ A ’ and ‘ B ’ both hold in this world, then ‘ C ’ also holds in this world’. For suppose that ‘ $(A \circ B) \rightarrow C$ ’ obtains in s . Then ‘ $A \rightarrow (B \rightarrow C)$ ’ is also true in s . If we hypothesise that there is a situation t such that ‘ A ’ is true in t and a situation u

³ To say that an inference form or a sequent is not valid is to say that at least one of its instances is invalid.

such that ' B ' holds in u , then we can infer that there is a situation v such that ' C ' is true in v .

Contrast this interpretation with our understanding of extensional conjunction. Suppose that ' $(A \wedge B) \rightarrow C$ ' holds in s . Then, if we postulate a *single situation* t in which ' A ' and ' B ' are both true, then we can infer that there is some situation u in which ' C ' is true.

The natural deduction rules for fusion are quite elegant. The introduction rule is very simple. It says:

$$(\circ I) \text{ From } A_\alpha \text{ and } B_\beta \text{ to infer } A \circ B_{\alpha \cup \beta}.$$

The introduction rule reinforces the relevant logicians' claim that fusion is really a sort of conjunction. We introduce fusions in proofs by taking together information in (perhaps) different situations. The elimination rule is also very clean. It is the following:

$$(\circ E) \text{ From } A \circ B_\alpha \text{ and } A \rightarrow (B \rightarrow C)_\beta, \text{ infer } C_{\alpha \cup \beta}.$$

The elimination rule tells us that having a fusion around allows us to perform a double modus ponens. This elimination rule very closely mirrors the meaning of fusion that we discussed above.

As we have seen, fusion has some very conjunction-like properties. It has an analogous connection with implication and negation to the one that extensional conjunction has with material implication and conjunction. It is also commutative, that is, the scheme,

$$(A \circ B) \rightarrow (B \circ A),$$

is valid. And it has the square increasing property: $A \rightarrow (A \circ A)$ is valid. But it also has some very non-conjunction-like properties. For example, idempotence fails for fusion. Thus,

$$(A \circ A) \rightarrow A$$

is not valid in R. As we saw earlier, simplification, viz.,

$$(A \circ B) \rightarrow B,$$

and,

$$(A \circ B) \rightarrow A,$$

are not valid in R. Idempotence is an instance of simplification, and it is interesting that even it is not valid in R.

9.5 Intensional disjunction

In addition to intensional conjunction, there is an intensional disjunction called *fission*. We can define fission – written $\oplus$ – in terms of fusion and negation, as,

$$A \oplus B =_{df} \sim(\sim A \circ \sim B)$$

or in terms of implication and negation, as,

$$\sim A \rightarrow B.$$

Our interest in fission is in its inferential role, which we will discuss in chapter 10 below. But here we will briefly discuss why logicians like to include it in the language of relevant logic.

The role of fission in sequents is the dual of that of fusion. Whereas fusion connects premises, fission can be taken to link conclusions together. The idea of having more than one conclusion might sound odd, but it adds a certain elegant symmetry and mathematical power to our proof theory. When Gentzen put forward the sequent calculus for classical logic, he had sequents with more than one conclusion. Gentzen interprets the sequent,

$$A_1, \dots, A_m \vdash B_1, \dots, B_n,$$

as saying the same thing as,

$$A_1 \wedge \dots \wedge A_m \vdash B_1 \vee \dots \vee B_n.$$

The use of multiple conclusions allows Gentzen to derive such classical principles as the law of excluded middle. It also allows him to do something that is mathematically very elegant. His sequent calculus for intuitionist logic is the same as the calculus for classical logic with the exception that in the system for intuitionist logic we are not allowed to have sequents with more than one conclusion. That makes for a very pretty relationship between classical and intuitionist logic.

In relevant logic, we can think of a sequent of the form,

$$A_1, \dots, A_m \vdash B_1, \dots, B_n,$$

as equivalent to,

$$A_1 \circ \dots \circ A_m \vdash B_1 \oplus \dots \oplus B_n.$$

Turning now to natural deduction, the easiest way to introduce fission in our system is by means of its relationship to implication. That is, we give it the following introduction rule:

Given a proof from $\sim A_{[k]}$ to B_α infer $A \oplus B_{\alpha - \{k\}}$, where k really occurs in α .

And, we have the elimination rule,

$$\text{From } A \oplus B_\alpha \text{ and } \sim A_\beta \text{ to infer } B_{\alpha \cup \beta}. \quad (9.1)$$

This elimination rule is sometimes called ‘Intensional Disjunctive Syllogism’ or ‘IDS’. In chapter 10 below, we will discuss IDS further.

9.6 The Scottish plan

Before we leave the issue of types of conjunction and disjunction, we should examine a very different interpretation of relevant logic due to Stephen Read (Read 1988).

Read urges us to accept intensional conjunction as a basic notion and use it to understand relevant implication. On his view, an implication, ‘ $A \rightarrow B$ ’, is to be understood as saying that it is not the case that $A \circ \sim B$. Thus the meaning of implication rides piggyback on the meaning of fusion. What does adopting fusion as a basic connective gain us? Read takes as a basic datum what he calls ‘the deduction equivalence’ which is a version of the combination of the deduction and converse deduction theorems. According to the deduction equivalence,

$$A_1 \text{ and } \dots \text{ and } A_n \text{ and } B \text{ entails } C$$

is equivalent to,

$$A_1 \text{ and } \dots \text{ and } A_n \text{ entails } B \rightarrow C.$$

What is interesting to us here is the use of ‘and’. If this equivalence does hold in normal English, then as relevant logicians we must hold that the ‘and’ here is not our normal extensional conjunction. For, as we have seen, taking ‘and’ to be extensional commits us to positive paradox. Read suggests that fusion is the right way of interpreting some uses of ‘and’ in natural language. Once we have accepted fusion (together with negation), we can define implication as,

$$A \rightarrow B =_{df} \sim(A \circ \sim B).$$

So far it looks like we have a tie. Read takes fusion as a primitive and I take implication as a primitive. There seems to be no real choice between the two views.

What decides whether either of these views is acceptable is its success in providing a coherent meaning for its primitive connective. The theory of situated inference is supposed to do that for implication in my theory. Read uses a proof theoretic account of meaning as the framework for his semantics of

fusion. Read's natural deduction system is different from ours (i.e. Anderson and Belnap's system) in its treatment of indices.⁴ Suppose that we have an inference with two hypotheses, $A_{[1]}$ and $B_{[2]}$. Then we can perform the operation of *bunching* these two hypotheses and get $A \circ B_{[1,2]}$. The index '[1,2]' is a multiset. It tells us that we have hypothesis 1 *and* hypothesis 2, in the sense of fusion. This explanation of the meaning of fusion in terms of fusing premises may seem circular, and in a sense it is. But this sort of circularity is not untoward in semantics. For consider the Tarskian truth condition for extensional conjunction:

' $A \wedge B$ ' is true if and only if ' A ' is true and ' B ' is true.

Here extensional 'and' is interpreted in terms of a metalinguistic 'and'. The case is very much the same in Read's proof theory. The fusion connective ($\circ$) is interpreted in terms of a metalinguistic fusion. What makes the Tarskian truth condition acceptable is that we have a pretheoretic understanding of extensional conjunction. Read thinks that we have a pretheoretic understanding of fusion as well. This pretheoretic understanding is demonstrated by our ability to group together premises in inferences.

It is difficult to compare Read's theory with my own. The framework in which I place my interpretation of relevant logic is the truth conditional theory of meaning. Read, however, agrees with the intuitionists that the meaning of the logical connectives is given by their roles in proofs ((Read 1988), chapter 9). Although I have accepted some of this position in my theory of situated inference, my theory remains thoroughly truth conditional. Read's proof-theoretic view of meaning allows him to say that whenever a logical particle has coherent rules of inference we can attribute to it a coherent meaning. From the point of view of the relational semantics, as we have said fusion seems like an oddity – a particle we would not expect to be expressed in natural language. Thus, I do not think that fusion can be taken to be a primitive notion, although it is useful for technical purposes in formulating a system of logic. This is not the place to decide between truth-theoretic and proof-theoretic semantics. I present Read's view because it is one of the very few other philosophical interpretations of relevant logic and the reader should be aware of it.

9.7 Consequence relations and theories

So far in this book we have been discussing inferences about the world made from the limited perspective of situations. But this isn't the only sort of inference that is useful to us. We sometimes want to use logical inference to help determine

⁴ Read's system is a development of the view of (Urquhart 1972) and similar ideas are developed in (Slaney 1990).

the content of *theories*. In chapter 4 above, in our discussion of paraconsistency, we briefly mentioned this aspect of logic.

Suppose, for example, that we have a theory of physics. We have a set of equations governing, say, electromagnetic fields. We can deduce from these equations certain properties that particles will have if they have certain initial positions and velocities in fields, changes in fields under certain conditions, and so on. The theory is not just determined by the equations themselves, but by inference rules, that include mathematics beyond those that are explicitly stated in the statement of the theory itself, and perhaps hidden in the background, logical rules. We will not merely be dealing with scientific theories here, but with the notion of a theory more generally.

A set of formulae Γ is a *theory* if and only if it meets the following two conditions:

- If $A_1, \dots, A_n$ are all in Γ and $A_1, \dots, A_n \vdash B$, then B is in Γ too (logical closure);
- If A and B are both in Γ , then so is $A \wedge B$ (conjunctive closure).

Following Kit Fine (Fine 1974), we can define a type of logical consequence that captures both logical and conjunctive closure. This is Fine's consequence relation:

$$\{A_1, \dots, A_m\} \vdash_{\wedge} \{B_1, \dots, B_n\} \text{ iff } \vdash_R (A_1 \wedge \dots \wedge A_m) \rightarrow (B_1 \vee \dots \vee B_n).$$

If the consequence set $\{B_1, \dots, B_n\}$ has only one member, say, B , then we may merely write ' $\{A_1, \dots, A_m\} \vdash_{\wedge} B$ '. And if the antecedent set $\{A_1, \dots, A_m\}$ has only one member, say, A , we may write ' $A \vdash_{\wedge} \{B_1, \dots, B_n\}$ '. This relation holds between sets (not multisets) of formulae. Where Γ or Δ (or both) are infinite sets, then we say that $\Gamma \vdash_{\wedge} \Delta$ if and only if there are some $A_1, \dots, A_m$ in Γ and some $B_1, \dots, B_n$ in Δ such that $\{A_1, \dots, A_m\} \vdash_{\wedge} \{B_1, \dots, B_n\}$ (where m and n are greater than or equal to 1). We can see that Fine's consequence relation is very much like Gentzen's turnstile for classical logic. The main difference between them is that our relation $\vdash_{\wedge}$ does not admit the deduction theorem.

We can use Fine's consequence relation to define a *theory*. First, we define a consequence operator, $Cn(\Gamma)$. Where Γ is a set of formulae,

$$Cn(\Gamma) = \{B : \Gamma \vdash_{\wedge} B\}.$$

A *theory* is a set of formulae, Γ , such that $\Gamma = Cn(\Gamma)$. In mathematical parlance, a theory is a fixed point for the consequence operator. It is an easy matter to show that the two definitions we have given for a theory are equivalent.

Soon we will see how useful theories can be.

9.8 Belief, denial and consequence

One sort of theory that will be of interest to us in chapters 10 and 11 below is the set of beliefs of an agent at a given time. Clearly, the beliefs of real agents are not typically closed under logical consequence. Our brains do not have the power to know all the consequences of our beliefs. But we will follow common practice in the literature on doxastic logic and belief representation⁵ in examining the beliefs of an ideal agent. Ideal agents do know all the consequences of their beliefs.

In addition to beliefs sets, we will examine *rejection sets*. An agent's rejection set is the set of statements that she denies. Recall what we said about denial in chapter 4 above. A denial cannot be reduced to any form of assertion. Rather, denial is a separate speech act from assertion. Corresponding to denial, there is the propositional attitude of rejection. The relationship between rejection and denial is the same as that between belief and assertion. The latter is the speech act that expresses the former.

Rejection sets are not theories. Rather, they are downwardly closed under $\vdash_{\wedge}$. A set of formulae, Γ , is downwardly closed under $\vdash_{\wedge}$ if and only if (1) if A is in Γ and $\{B\} \vdash A$ then B is also in Γ and (2) if A is in Γ and B is in Γ then $A \vee B$ is in Γ as well.⁶ Or, more concisely, where $Do(\Gamma)$ is the downward closure of Γ ,

$$Do(\Gamma) = \{B : B \vdash_{\wedge} \Gamma\}.$$

The rationale for the downward closure of rejection sets is quite simple. If you reject a proposition, you should also reject anything that entails it. More generally, if you reject a set of propositions, you should reject anything that entails a disjunction of elements from that set.

Although assertion and denial are separate speech acts, they do have an interesting relationship to one another. First, given the semantics for rejection that an ideal agent's acceptance set and her rejection set should be exclusive, or formally,

$$Acc(a) \cap Rej(a) = \emptyset,$$

where a is an ideal agent. This we will call the postulate of *pragmatic consistency*. Note that pragmatic consistency and normal consistency do not imply one another. One might have a consistent acceptance set and yet a rejection set that overlaps with it and one might have an inconsistent acceptance set and yet have a rejection set that does not overlap with it.

⁵ For some reason, this field is usually called *knowledge representation*, when it is belief, not knowledge, that is the real object of study.

⁶ For readers with some algebraic logic: a theory is a filter and a rejection set is an ideal on the Lindenbaum algebra of the logic.

Moreover, it would seem that there are some pragmatic versions of standard rules of inference that should govern an ideal agent. A good example is modus tollens. That is, if for an ideal agent a ,

$$(A \rightarrow B) \in \text{Acc}(a)$$

and,

$$B \in \text{Rej}(a),$$

then,

$$A \in \text{Rej}(a).$$

In chapter 10 we will discuss other pragmatic rules.

9.9 Consistency and strong consistency

We will need a few more concepts for the chapters that follow. The first of these is a standard logical concept – consistency. We have already been using this term in its usual sense of ‘negation consistency’. A theory is consistent if and only if for no formula does it contain both the formula and its negation. Consistency is a useful concept in the philosophy of relevant logic, but there are other similar conceptions that are useful as well.

One important concept is that of *strong consistency*. Although this concept is often used, neither this name nor any other name is usually given to it. In order to understand strong consistency, we first need to know what a prime theory is. A theory Γ is *prime* if and only if for every disjunction $A \vee B$ in Γ either A is in Γ or B is in Γ .

A theory is *strongly consistent* if and only if it is characterised by the set of its prime, consistent extensions. A theory Δ is an extension of a theory Γ if and only if Γ is a subset of Δ . A theory Γ is characterised by a set of theories if and only if Γ is just the intersection of that set, that is, Γ includes all and only those formulae that are in all of the theories in that set.

The notion of strong consistency has been used to prove that Wilhelm Ackermann’s rule gamma (γ) is admissible in various relevance logics. This technique of proving the admissibility of γ originally was developed by Bob Meyer and has been adapted by him and other logicians to treat a wide range of systems. The rule gamma is the following:

$$\frac{\begin{array}{c} \vdash \sim A \vee B \\ \vdash A \end{array}}{\vdash B}$$

Ackermann’s formulation of his theory of entailment (his system Π') included γ . Anderson and Belnap object to the inclusion of γ . They point out that

we cannot include disjunctive syllogism (the rule that allows us to infer from $\sim A \vee B$ and A to B) in our natural deduction system ((Anderson *et al.* 1992), §45.3). As we shall see in chapter 10 below, the addition of this rule (or any variant of it) leads to the acceptance of *ex falso quodlibet*. The problem, as they see it, is that if we can use a rule to reason about theorems (as γ is supposed to do), then we can use it to reason about anything at all, that is, we can use it to reason about hypotheses. For rules of deductive reasoning are supposed to be perfectly general. Thus, they formulate their logics (E and R) without the rule γ and only using rules of implication elimination and conjunction introduction, which do have corresponding rules in the natural deduction system.

But the problem then arose as to whether Anderson and Belnap's logic E, that does not contain γ , has the same set of theorems as Ackermann's logic. Meyer and Dunn first proved that it does have the same theorems (Meyer and Dunn, 1969), but the proof that interests us here is one that Meyer did on his own using the property of strong consistency.

In order to prove that γ holds in a logic L , we show that it is strongly consistent. We take the set of theorems of L and show that this theory is characterised by its prime, consistent extensions. We then suppose that A and $\sim A \vee B$ are theorems of L . Then we know that both of these formulae are in every prime, consistent extension of the theorems of L . Since each extension is consistent, $\sim A$ is not in it. But $\sim A \vee B$ is in it and it is prime. Therefore, for each prime, consistent extension of the set of theorems of L , B must be in that extension. Since the set of theorems is characterised by the set of such extensions, B is a theorem of L , thus completing the proof of the admissibility of γ .

Note that our proof uses a version of disjunctive syllogism. That is, given Γ , a prime, consistent extension of L , we infer from,

$$\sim A \notin \Gamma$$

and,

$$\sim A \in \Gamma \text{ or } B \in \Gamma$$

to,

$$B \in \Gamma.$$

This is clearly an instance of disjunctive syllogism. In the next chapter we will discuss what is wrong with disjunctive syllogism and when and how we can justify using it.

10 Disjunctive syllogism

10.1 The problem of disjunctive syllogism

We ended the previous chapter by saying that there is a problem with disjunctive syllogism. In this chapter we will look in depth at that problem and propose some ways of avoiding it.

Two versions of the standard rule are the following:

(DS) From $A \vee B_\alpha$ and $\sim A_\alpha$ to infer B_α .

(SDS) From $A \vee B_\alpha$ and $\sim A_\beta$ to infer $B_{\alpha \cup \beta}$.

Disjunctive Syllogism (DS) is an instance of Strong Disjunctive Syllogism (SDS) – just set $\beta = \alpha$. Unfortunately, if we add either of these rules to the natural deduction system, we no longer have a relevant logic. Instead, we have classical logic.

A proof that we get classical logic can be abstracted from the argument that we discussed in the introductory chapter above that C. I. Lewis used in support of *ex falso quodlibet*. Rewritten in the natural deduction system, Lewis' argument is the following:

- | | |
|--------------------------------------|----------------------|
| 1. $A \wedge \sim A_1$ | <i>hyp</i> |
| 2. $A_{\{1\}}$ | 1, $\wedge E$ |
| 3. $A \vee B_{\{1\}}$ | 2, $\vee I$ |
| 4. $\sim A_{\{1\}}$ | 1, $\wedge E$ |
| 5. $B_{\{1\}}$ | 3,4, DS |
| 6. $(A \wedge \sim A) \rightarrow B$ | 1–5, $\rightarrow I$ |

Note that this argument goes through with either form of DS. Lewis' argument shows that we cannot add disjunction elimination to our natural deduction system without also obtaining *ex falso quodlibet*.

The fact that disjunctive syllogism cannot be added to relevant logic is a problem. The original point of introducing relevant logic was to provide an intuitive characterisation of deductive inference. But we use disjunctive syllogism all the time. When we make inferences while reading a murder mystery

we use it. We eliminate suspects to decide who the killer is. When we look for things we have misplaced we use it ('It is either in the bedroom or the kitchen. It isn't in the kitchen. Therefore it must be in the bedroom'). In short, disjunctive syllogism would seem to be one of our key deductive tools. So it looks like we have a problem.

10.2 Garfield's dog

One way of dealing with the problem is to deny that we really have a problem. That is, we can deny that disjunctive syllogism is a good principle of reasoning after all.

Jay Garfield takes this route (Garfield 1990). Garfield argues that we should avoid using disjunctive syllogism under conditions of 'epistemic hostility'. An environment is epistemically hostile when there is important misleading information available in it.

Garfield's argument is quite simple. He claims that under conditions of epistemic hostility the use of disjunctive syllogism can leave us with unwarranted and dangerous beliefs. He says:

Suppose that in these unfortunate but all too common circumstances you come to believe on the misleading information of a reliable source (*A*) Albuquerque is the capital of Arizona. Under the spell of an evil classical logician you freely disjoin (*B*) Belnap is a classical logician. Since *A* is justified, so is $A \vee B$. Now, suppose that a bit later your geographical source corrects himself, and you come to believe that $\sim A$. Now $\sim A$ and $A \vee B$ are both in your belief set, you have no positive reason to reject $A \vee B$, you are still classical, so you conclude *B*. *B*, of course, is manifestly false. What's more, from a suitably distant perspective (ours) you have no real reason to believe it. What went wrong? The answer is plain: you used classical disjunction rules. (Ibid., p. 104)

This argument follows the form of C. I. Lewis' argument for *ex falso quodlibet* given in section 10.1 above. It depends on the use of disjunction introduction and then disjunctive syllogism. The difference is that instead of using a contradictory premise, Garfield's argument conjures up a situation in which the agent changes his or her mind. People generally do not keep track of the justificatory sources of their beliefs and so even when the justification is removed for an individual belief, in many cases, the agent will retain the belief without any justification.

There is something right about the conclusion that Garfield draws. Under conditions of epistemic hostility we can run into difficulty by using disjunctive syllogism. But there is also something wrong with Garfield's argument. Disjunctive syllogism is not unique with regard to its capacity to lead us astray under conditions of epistemic hostility. Let's try another case. Suppose that you have heard from a usually reliable source that it is a law of nature that substances

always contract when cooled ($\forall x(K(x) \rightarrow C(x))$, where the domain is the set of substances). You instantiate this to a belief that water will always contract when cooled ($K(w) \rightarrow C(w)$) and you believe that this has the status of a law of nature. You put a bottle of water in the freezer and believe that the water in the bottle's being cooled implies that it will contract. Later you find out that not all substances contract when cooled. Some crystallise when frozen and expand. Thus you reject the belief that $\forall x(K(x) \rightarrow C(x))$. But you haven't discovered that water is one of those substances that crystallises and expands. So you reject your original information, but you still believe that the water in the bottle in your freezer has contracted. Meanwhile, unknown to you, the bottle has burst outwards. Once again, you have drawn a false conclusion and it looks like you no longer have good justification for it.

What went wrong this time? The two logical rules that you have used are universal instantiation and modus ponens. Some relevance logicians reject modus ponens, but blaming it would be beside the point here. We can cook up similar examples using almost any rule of inference. It looks like we can get into trouble under conditions of epistemic hostility if we make any substantive inferences on the basis of what we think we discover. Perhaps we should not look to cutting rules out of our logic but rather to other means of risk management to deal with epistemic hostility.

In sum, I don't think that Garfield's argument against disjunctive syllogism works. If it did, it could be used for any purported logical rule, except perhaps that we can derive a proposition from itself. Thus, the argument at best can be taken to be an argument against using logic in epistemically hostile situations. But it cannot be used to show that individual rules are not part of logic.

10.3 When don't we want to use disjunctive syllogism?

We will be able to bring the problem into sharper focus if we look at cases in which we think that disjunctive syllogism is not to be used.

One domain in which we cannot always use disjunctive syllogism is in reasoning about fictions. Fictions sometimes contain contradictions and, as we have seen, if we are allowed unconditional application of disjunctive syllogism, then we can derive that any proposition holds in such fictions.

Thus, it would seem that we have a class of rather straightforward counterexamples to disjunctive syllogism. But, in fact the topic of inconsistent fictions is rather more complicated than this. It has been suggested by some authors (see, e.g., (Lewis 1984)) that we can deal with inconsistent fictions within a possible world framework. On this view, we think of a fiction that includes a contradiction as determining two classes of worlds. One class contains only

worlds that make one of the contradictory conjuncts true and the other class contains worlds that make the other conjunct true. What is true in the fiction is what is true in every world in the one class *or* is true throughout the other class.

This 'non-adjunctive approach' should be accepted for some fictions, even within the relevant-situational framework that we are using here. Suppose, for example, that we have a fiction in which we are told at one point that at a specific time Ms Jones has exactly two children. Later in the book we are told that she has exactly one child at that same time. Should we infer from this that $2 = 1$? Together with a bit of elementary arithmetic, this leads quickly to disaster. For if we subtract one from each side of the equation, we obtain $0 = 1$, and from this we can prove that $n = m$ for any natural numbers n and m . We don't want to be able to do that. It is easier to consider the fiction as determining two classes of situations, one which contains situations which make it true that she has two children and the other which contains situations which make it true that she has one child.

Disjunctive syllogism isn't valid for fictions on this view and neither are modus ponens, conjunction introduction, or a host of other rules. But it provides little aid or comfort for our cause. What we need is to claim that there are fictions that require individual situations to fail to be closed under disjunctive syllogism. Then we will be justified in claiming that disjunctive syllogism is intuitively invalid.

It would seem, however, that there can be fictions that suit our needs. Consider for example a complex tale in which a minor character dies. We are told at one point that he leaves his small estate to Jones instead of Smith. For this reason, Smith becomes jealous of Jones. Much later in the book, we are told that the very same estate is left (at exactly the same time) to Smith, and not Jones. At a yet later point in the book, Smith is said to be both jealous of Jones (this is carried over from the original slighting of Smith in the will) and rich enough to do something about it (because he inherited the will). He hires a contract killer to kill Jones. The author, without realising it, has written an inconsistent story. And he uses the inconsistency without realising it.

If we split up the situations that are used to provide a semantics for the story into two classes according to the non-adjunctive approach, then we will ruin the narrative structure of the story. The fact that Smith is jealous is explained by his being excluded from the will. The fact that he has enough money to hire a contract killer is to be explained by the fact that he was not excluded from the will. Moreover, the fact that he did hire a contract killer is to be explained, in part, by the fact that Smith was jealous of Jones. It seems that we have to leave these strands intertwined in order to retain the narrative integrity of the text.

Now let's return to disjunctive syllogism. It seems that we need situations in which it is both true that

Jones inherits the estate

and

Jones does not inherit the estate.

From this we would hardly want to infer that Jones hates ice cream. But we could do so if we are allowed disjunctive syllogism, for we can infer from Jones inherits the estate that either he inherits the estate or he hates ice cream. Using disjunctive syllogism and the fact that Jones does not inherit the estate, we could infer that Jones hates ice cream. So it seems that we should reject disjunctive syllogism as a logical rule of inference.

10.4 Read's way out

But we can't reject disjunctive syllogism altogether. It seems to be a rule used by some people when they are reasoning well. How do we explain this?

One explanation is championed by Stephen Read. Read argues what looks like DS in natural language is in fact an argument rather closely related to DS (or, rather, to SDS), but is valid in relevant logic. This valid argument is the intensional disjunctive syllogism (IDS). IDS is the following:

From $A \oplus B_\alpha$ and $\sim A_\beta$ to infer $B_{\alpha \cup \beta}$,

where ' $\oplus$ ' is the connective fission, discussed in the previous chapter. As we have seen, in our logic, fission can be defined as follows:

$$A \oplus B =_{df} \sim A \rightarrow B.$$

Given this definition, one can see that IDS is really just a version of implication elimination. For it allows us to infer from $\sim A \rightarrow B_\alpha$ and $\sim A_\beta$ to $B_{\alpha \cup \beta}$.

The problem is trying to show that what we usually take to be instances of DS in natural language are in fact instances of IDS. (Read 1983) uses the following example taken and modified from (Burgess 1981). Suppose there is a card game which uses three cards, call them ' A ', ' B ', and ' C '. Two cards are placed face down on a table and the other is in the hand of my 'truthful but uncooperative partner' ((Read 1983), p. 474). I am supposed to guess which are the two cards on the table by asking, say, whether one is A and the other B . My partner may answer either 'no' or 'maybe'. Suppose that I guess first that B and C are on the table and then that A and C are on the table and I am told 'no' both times. I thus conclude that C is not on the table, hence that A and B are.

We can formalise this argument as:

$$\begin{array}{l} \sim(B \wedge C) \\ \sim(A \wedge C) \\ \hline \therefore A \wedge B \end{array}$$

As Read points out, this argument is an enthymeme. It requires an extra premise before we can count it as valid, even in classical logic. Clearly, the missing premise is

$$(A \wedge B) \text{ or } (A \wedge C) \text{ or } (B \wedge C).$$

Read now asks what sort of disjunction ‘or’ represents here. He answers that it is fission, not extensional disjunction. Read says:

They are intensional. For the truth of the three-part disjunction does not depend solely on the truth of some single disjunct, e.g., that as a matter of fact *A* and *B* are the cards on the table. The other disjuncts are inferentially connected to $(A \wedge B)$ in a way in which those in

$$(A \wedge B) \text{ or Bach wrote the Coffee Cantata or the Van Allen belt is doughnut shaped}$$

are not. The truth of [the above statement] arises simply from the fact that *A* and *B* are on the table; whereas that of (the missing premise) depends on the fact that there are only three possibilities (since there are only three cards), so that *if* two of them are not realised, then the third must be. The disjunction has the force of a conditional. ((Read1983), p. 475)

Read’s test to determine whether a disjunction is intensional is to ask whether we can infer a conditional from it. If we can, it is intensional.

Read claims that the missing premise passes the test, and so represents the full argument as:

$$\begin{array}{l} (A \wedge B) \oplus (A \wedge C) \oplus (B \wedge C) \\ \sim(B \wedge C) \\ \sim(A \wedge C) \\ \hline \therefore A \wedge B \end{array}$$

And this is just a double application of IDS. Hence it is a valid argument from a relevant point of view.

I think Read is almost right. His test for intensional disjunction is not strong enough. The mere fact that a conditional obtains does not indicate that the corresponding intensional disjunction holds. In order to work, Read does not need to show that a conditional holds, but rather that a relevant implication obtains.

I do not think, however, that a relevant implication does obtain here. The extensional disjunction ‘ $(A \wedge B) \vee (A \wedge C) \vee (B \wedge C)$ ’ does hold, but all

we can infer from this is that an indicative conditional obtains. As we saw in chapter 6 above, this inference is not a deductive inference but a reasonable inference never the less.

Let's consider the example again. Suppose that s is a situation in which the game is being played, and that its rules are being obeyed. I maintain that in s , the conditional

$$\sim(B \wedge C) \Rightarrow ((A \wedge C) \vee (A \wedge B))$$

would likely be true.

It would seem in s , that it is legitimate for us only to consider circumstances in which the rules of the game are obeyed. Thus, in the pairs of situations that we will use to evaluate the conditional above the following disjunction holds:

$$(A \wedge B) \vee ((A \wedge C) \vee (B \wedge C)).$$

Moreover, since the rules of the game are consistent, it would seem that we would look only at possible circumstances to evaluate the conditional. And so, as we saw in chapter 6 above, the conditional holds.

So now we have an inference of the following form:

$$\frac{\begin{array}{c} \sim(B \wedge C) \Rightarrow ((A \wedge C) \vee (A \wedge B))_{\{s\}} \\ \sim(B \wedge C)_{\{s\}} \end{array}}{\therefore (A \wedge C) \vee (A \wedge B)_{\{s\}}}$$

Thus, we have an instance of modus ponens.

But we are not done yet. We have inferred that ' $(A \wedge C) \vee (A \wedge B)$ ' is true in s and what we need is some way of eliminating ' $A \wedge C$ '. This is done by the same method. Now we have in mind a set of circumstances in which the rules of our game are obeyed and ' $(A \wedge C) \vee (A \wedge B)$ ' is also true. So, we have

$$\sim(A \wedge C) \Rightarrow (A \wedge B)$$

in s . Thus, when we are told that $\sim(A \wedge C)$, we use modus ponens to infer that $(A \wedge B)$.

Clearly, this reconstruction of the reasoning in this example is highly idealised. What it does not pretend to be is a description of the psychological processes of someone playing the game. People do use disjunctive syllogism without going through all these steps. What we are doing instead is showing how we can reconstruct the reasoning in a way to make it clearly reasonable. Having done that, we vindicate the use of disjunctive syllogism in contexts like this.

10.5 Consistency as a premise

Let's think about disjunctive syllogism in another way. As we saw in chapter 9 above, when we have a theory that is strongly consistent that theory is closed under disjunctive syllogism. So why not allow disjunctive syllogism when we add the premises that we have a strongly consistent theory?

Chris Mortensen advocates such a position (?). He thinks that we should accept the following entailment. Where, ' $Con(x)$ ' means that x is consistent and ' $Pr(x)$ ' means that x is prime, and we add quasi-quotation marks (' $\ulcorner$ ' and ' $\urcorner$ ') and the binary relation ' $\in$ ' ('belongs to') to our language,

$$(Con(Th) \wedge Pr(Th) \wedge \ulcorner A \vee B \in Th \urcorner \wedge \ulcorner \sim A \urcorner \in Th) \rightarrow \ulcorner B \urcorner \in Th.$$

As Mortensen points out, adding this scheme to a logic like E or R does not force the logic to accept ex falso quodlibet. Ex falso quodlibet is still invalid when we add this scheme ((Mortensen 1986), p. 197).

There is something right about this idea. When some relevant logicians at least prove results about relevant logics, such as a proof that the rule gamma is admissible in a particular system, when we prove that a theory is normal we conclude that it is closed under disjunctive syllogism. This sort of reasoning, it might seem, could be cashed out as an application of Mortensen's scheme given above.

Unfortunately we must reject Mortensen's view. We reject it, not because there is something intrinsically wrong with it, but because it cannot fit into our semantical framework. Consider the truth condition for the statement ' $Con(Th)$ '. It would seem reasonable to set this as

$$\text{'Con(Th)' is true at } s \text{ if and only if for no formula 'A', is it true in } s \text{ that } \ulcorner A \urcorner \in Th \text{ and } \ulcorner \sim A \urcorner \in Th.$$

The problem here is that we might have a situation s which satisfies the right-hand side of this biconditional but a situation s' , which contains s , which does not satisfy it. Thus, $Con(Th)$ will be true at a situation s , but not at all situations that contain s . In other words, $Con(Th)$ will not express a proposition.

Still there is something right about Mortensen's intuitions. If we say that a theory is consistent and prime, we should be able to use disjunctive syllogism to make inferences about its contents.

Like Read's view, we can make Mortensen's view more natural by thinking in terms of indicative conditionals rather than in terms of implication. Again, we use our conditional from chapter 6 above for this purpose. So, let's reformulate Mortensen's argument.

Primeness, in fact, is a slightly stronger property than the one we need. What we will require here is that a theory be characterised by its consistent, prime

extensions, that is, a theory that is strongly consistent. Suppose that Th is a theory that is strongly consistent. The conditional that will be of interest to us here is:

$$(\ulcorner \sim A \urcorner) \in Th \Rightarrow \ulcorner B \urcorner \in Th.$$

If, when we evaluate this conditional we consider only possible circumstances in which Th is strongly consistent and in which

$$(\ulcorner A \vee B \urcorner) \in Th,$$

it is reasonable to infer this conditional.

Thus we have a reconstruction of disjunctive syllogism. From,

Th is strongly consistent,

and,

$$(\ulcorner A \vee B \urcorner) \in Th,$$

we can *reasonably* infer,

$$(\ulcorner \sim A \urcorner) \in T \Rightarrow \ulcorner B \urcorner \in Th.$$

From this and the premise,

$$(\ulcorner \sim A \urcorner) \in Th,$$

by modus ponens for the conditional, we can derive,

$$\ulcorner B \urcorner \in Th,$$

which is what we want.

10.6 Denial and disjunctive syllogism

We have already seen that we can save disjunctive syllogism by using our theory of the conditional. Here is a different approach, which I had originally presented in (Mares 2000) (and owes a good deal to (Priest 1986)).

Recall from chapter 4 above the distinction between negation and denial. Denial is a speech act. To deny a proposition is to hold that it fails to obtain in the context in which the denial is made. Corresponding to denial is the propositional attitude of rejection. The relationship between denial and rejection is the same as that which holds between assertion and belief.

We can formulate a pragmatic version of disjunctive syllogism (PDS) using denial (or rejection). If one accepts,

$$A \vee B$$

but rejects,

A ,

then she should also accept,

B .

This rule seems reasonable. For anyone¹ who understands the meaning of extensional disjunction realises that accepting a disjunction commits her to accepting that at least one of the disjuncts is true. By denying that one disjunct is true, then, she tacitly (at least) commits herself to accepting the other disjunct.

For PDS to be of any use, we also need to accept the following inference, that we will call PDS'.² Suppose that we reject,

$A \wedge B$

and we accept,

A ,

then we should also reject,

B .

For if the rejection of the conjunction is accurate and so is the acceptance of A , then the other conjunct cannot also be true (or else the conjunction would be true as well). Thus, we should reject it as well.

I also claim that we should reject contradictions. As we saw in chapter 5 above there are no true contradictions. We should reject contradictions because the view that there are no true contradictions is very deeply entrenched and, by and large, we should be conservative in how we abandon our views. But there are times, perhaps, in which contradictions seem forced on us, by our best scientific and mathematical theories, and so on (see chapter 5). Thus, we should reject contradictions in most cases so much so that we should consider it to be a default assumption.

How does all of this bear on the problem of disjunctive syllogism? Suppose that an agent accepts a disjunction,

$A \vee B$

and a negation,

$\sim A$.

¹ anyone in a possible situation at least.

² This inference is mis-stated in (Mares 2000) on page 509. The statement there is marred by a typo, but a very misleading typo.

By default she should also reject the contradiction,

$$A \wedge \sim A.$$

By PDS', the agent may then reject,

$$A.$$

By PDS and the fact that she accepts $A \vee B$, she may then accept,

$$B.$$

Through this roundabout route, we justify the use of disjunctive syllogism.

Thus, there are at least three ways in which a relevant logician can justify using disjunctive syllogism. I do not know whether there are other reasonable justifications, but for our purposes it does not matter. What does matter is that we have shown that the relevant logician is not committed to holding that this intuitive principle of reasoning is not reasonable.

10.7 Vindicating classical mathematical reasoning

Mathematicians use classical logic, at least many of them do. Their reasoning is usually taken to be good. And I think that it is too. We have all the materials in our theory to vindicate the use of classical logic in mathematics.

Mathematicians talk in terms of theories. There is set theory, number theory, the theory of the reals, the theory of the hyperreals, there are various geometries, theories of various groups, and so on. Suppose that T is some mathematical theory that is thought to be consistent. On our reading, the standard use of 'consistent' means that it is thought to be characterised by its normal extensions.

The claim that it is consistent carries with it the pragmatic force that we should, all things being equal, reject the statement that there are contradictions in T . Thus, $\exists p(\ulcorner p \urcorner \in T \wedge \ulcorner \sim p \urcorner \in T)$ is in our rejection set. Using the downward closure of rejection sets, for every statement ' A ', $\ulcorner A \urcorner \in T \wedge \ulcorner \sim A \urcorner \in T$ is at least tacitly in our rejection set. Using PDS and PDS', as we have detailed in the previous section, we can then treat T as closed under disjunctive syllogism.

Disjunctive syllogism is not all that there is to classical reasoning. We need, for example, a set of axioms for classical logic. Mathematicians tend to help themselves to those theorems of classical logic that they know. For example, proofs by cases often appeal to the law of excluded middle. Thus, we can say that mathematical theories at least often contain a set of axioms for classical logic (even if they are not explicitly acknowledged). Now we have a vindication of classical reasoning about some theories.

Sometimes, however, we want to use these theories to talk about reality too. To do so, we can use the following logical truth.³ For all statements ‘ A ’ it is a logical truth that,

$$(Tr(T) \wedge \ulcorner A \urcorner \in T) \rightarrow A.$$

In words, if a theory is true and a statement is in that theory, then that statement holds. Thus, if a theory is consistent (in our strong sense), then we can use disjunctive syllogism to make inferences about its contents. If that theory is true, we can infer that its contents are also true.

I began this book with a complaint that classical mathematical reasoning does not capture the notion of a proof. Now we have vindicated classical mathematical reasoning. Am I contradicting myself? I don’t think so. Consider again the problematic ‘proof’ from chapter 1 above:

The sky is blue.

∴ There is no integer n greater than or equal to 3 such that
for any non-zero integers x, y, z , $x^n = y^n + z^n$.

This ‘proof’ relies on the validity of the argument form,

$$\frac{A}{\therefore B}$$

where B is a necessary truth according to the theory that we assume (in this case the theory of arithmetic). In terms of our natural deduction system, we can construe this as follows:

1. $\ulcorner B \urcorner \in \mathbb{N}_{\{1\}}$ *hyp*
2. $A_{\{2\}}$ *hyp*
3. $\ulcorner B \urcorner \in \mathbb{N}_{\{1\}}$ 1, *reit*.

where $\mathbb{N}$ is the theory of arithmetic. This is not a good relevant deduction, since the index for the second premise does not appear in the subscript for the conclusion. What we do accept, however, is the rather harmless inference form

1. $\ulcorner A \wedge B \urcorner \in \mathbb{N}_{\{1\}}$ *hyp*
2. $\ulcorner B \urcorner \in \mathbb{N}_{\{3\}}$ 1, (and a complicated version of $\wedge E$)

In other words, if we take the premises to be conjunctive, in the sense of chapter 9 above, then the argument is valid in our semantics. If we take the premises to be intensionally connected, however, we must reject it.

Nor do we accept *ex falso quodlibet*. As soon as one suggests that a contradiction might hold in a theory, the rejection of all contradictions being true

³ I don’t want to get into a discussion of truth, the liar paradox, and so on, right here. Let it suffice to say that this implication is true at least for most theories T (ones that do not contain any self-reference, say).

of that theory has to be held in suspension. Hypothesising that a contradiction holds in a theory forces us seriously to consider what would be the case if there is a contradiction in the theory. We then cannot use PDS and PDS' without restriction to justify the use of disjunctive syllogism. Thus, if a contradiction is hypothesised to hold in a theory, our vindication of classical reasoning breaks down. And then we certainly are not entitled to use *ex falso quodlibet*.

So our *vindication of classical reasoning* does not really vindicate all of classical reasoning. It does not let us infer arbitrary truths (even arbitrary necessary truths) from any old proposition and it does not allow us to use *ex falso quodlibet*. What does it vindicate? It vindicates the use of a Hilbert-style deduction system where theorems are deduced from a strongly consistent set of axioms and rules (and which is such that we have good reason to believe that it is strongly consistent). It does not vindicate the use of a natural deduction system for classical logic which allows arbitrary hypotheses (such as inconsistent hypotheses).

10.8 The problem of a metalanguage

Now we can see quite clearly how we can use a classical metalanguage in this book. I have assumed that the metatheory of this book is consistent. I think I am justified in so doing. The Routley–Meyer semantics has been around for thirty years and has been well used by many logicians. No contradiction has been produced in it (although it has been used to model contradictions). Similarly, the other theories I employ, such as Aczel's non-well-founded set theory have been well examined, and Aczel has a proof of the relative consistency of his theory with standard Zermelo–Frankel set theory, with the axiom of foundation (ZF). ZF in turn has been extremely widely used without mathematicians finding any contradictions.

Thus, we have good inductive evidence for the belief that the theories we are using are consistent, even in the strong sense that we mean here. We thus have good justification for employing the methods outlined above to justify the closure of the theory of this book under disjunctive syllogism. Thus, our use of the material conditional and the material biconditional in stating truth conditions, and so on, is reasonable. We can use detachment (*modus ponens*) for the material conditional and bi-conditional. And that is really all we require to vindicate our use of a classical metatheory for relevant logic.

11 Putting relevant logic to work

11.1 Doing things with logic

We have already discussed various jobs that relevant logic can do. It provides us with a theory of situated inference, a theory of implication, and the basis of a theory of conditionals. In this chapter we will use those theories to do other work that is of interest to philosophers, mathematicians and computer scientists.

11.2 Dyadic deontic logic

In chapter 6 above, we briefly discussed deontic logic. There we discussed the operator, ‘it ought to be that’. There are also logics that contain a dyadic deontic operator, ‘it ought to be that ... on the condition that’. This is usually written in formal notation as ‘ $O(-/-)$ ’. The formula, ‘ $O(B/A)$ ’, is read ‘It ought to be that B on the condition that A ’. In this section, I will argue that the dyadic deontic operator should be taken to be some sort of relevant counterfactual.

The link between counterfactual conditionals and dyadic deontic operators is quite striking. Like a counterfactual, the connection in a dyadic deontic statement between antecedent and consequence is defeasible. For instance, suppose that my friend Kevin asks to borrow a length of rope from me. I know that, on the condition that Kevin is a friend and that he asks to borrow a rope, I should lend it to him. But suppose that I then discover good evidence that Kevin is extremely depressed, so much so that I fear that he is suicidal. Then my obligation to lend Kevin the rope is annulled, since I fear that he may hang himself with it.

Before we get to the relevance properties of the dyadic operator, let us look at what is perhaps the most popular theory. This is the one developed by Bengt Hansson (Hansson 1969), Bas van Fraassen (van Fraassen 1972) and David Lewis ((Lewis 1973), §5.1). This theory gives a possible world semantics for the operator – $O(B/A)$ is true at w if and only if in all the best worlds closest to w in which A obtains, B obtains as well. This semantics is quite intuitive. We ought to do B on the condition that A if and only if in the best worlds in which A is true B happens as well.

The van Fraassen–Lewis theory, however, has a particularly odd consequence. It can be proved that, in every possible world $O(A/A)$. For example, on this view it is true that it ought to be the case that the Manson murders were committed on the condition that the Manson murders were committed. This seems extremely strange. The violation of a moral rule does not thereby make that violation an obligation.

A better view, in my opinion, is that of Brian Chellas ((Chellas 1974) and (Chellas 1980), §10.2). On Chellas' view, the formula $O(B/A)$ is defined as,

$$A \Box \rightarrow OB,$$

where O here is the standard deontic operator, 'it ought to be the case that'. We will discuss the nature of the counterfactual in this definition later, but it is immediately obvious that now we do not make $O(A/A)$ valid for every formula A .

But Chellas' view still has difficulties. Like the van Fraassen–Lewis view, it is based on possible world semantics. Suppose for example that in every closest world in which Sean goes to the shops, it ought to be that Susan should not murder anyone. Thus, on Chellas' view, we get,

Susan should not murder anyone on the condition that
Sean goes to the shops.

This seems very strange indeed. Surely, there should be more of a connection between Sean's going shopping and Susan's refraining from murdering than is required by this semantics. What is needed is an injection of relevance.

One way of dealing with this problem is to use a relevant counterfactual instead of Chellas' classically based counterfactual. For example, we could use the theory of counterfactuals of chapter 8 above. To use this theory, we need some idea of how to specify a set of base situations for a given counterfactual. In each context, for each dyadic deontic claim, there are a set of default conditions that we take for granted. These default conditions specify the salient base situations. Take, for example, the earlier example of Kevin's asking me to borrow a rope. I associate with the claim 'On the condition that Kevin asks me to borrow a rope and he is a friend, I should lend it to him' the default conditions that Kevin is sane and in reasonably good spirits. At least I should associate these propositions with the claim if I do not have any evidence to the contrary, since most people I know (including Kevin) are sane and in reasonably good spirits the vast majority of the time. In these base situations, the conditional 'If Kevin asks to borrow a rope and he is my friend, I should lend it to him' comes out true. Thus, the dyadic deontic claim comes out true. But, if we shift the context of utterance to one in which I now have good reason to believe that Kevin is seriously considering hanging himself, then we will no longer choose

those sorts of base situations and the dyadic deontic claim itself will turn out to be false.¹

There is another treatment of dyadic deontic logic using relevant logic. This is due to Lou Goble (Goble 1999). Goble adds a relation T to the Routley–Meyer semantics for relevant logic. The relation T holds between pairs of situations and formulae. He gives the following truth condition for the dyadic deontic connective:

‘ $O(B/A)$ ’ is true at s if and only if for all t such $TsAt$, B is true at t .

I have no objections to Goble’s view. The relation T does, however, need a philosophical interpretation. Perhaps when that interpretation is provided it will seem like the correct approach to the problem. My object in this section is to provide an argument that we need a relevant treatment of dyadic deontic logic, not necessarily that we should use my theory of counterfactuals as a basis for it.

11.3 Essential predication and relevant predication

Another use of modal relevant logic is in the formulation of a theory of essential predication. The theory I have in mind is due to Dunn (?). In those papers Dunn sets out a theory of ‘relevant predication’ and treats essential predication as a form (or, rather, three forms) of necessary relevant predication.

Consider the difference between the following two statements:

Ramsey is such that he is furry.

Ramsey is such that Socrates is wise.

The first of this pair seems reasonable in a way that the second does not. Ramsey’s being furry has something to do with him. The predicate in the first statement really is predicated of the subject. In the second sentence, the subject and predicate seem accidentally related. Socrates’ wisdom has nothing to do with Ramsey. Dunn claims that the predication in the second statement is *relevant* and the first statement is a case of irrelevant predication.

Dunn holds that an individual i has a property P relevantly if and only if

$$\forall x(x = i \rightarrow P(x)).$$

The idea here is quite simple. i has P relevantly if and only if being identical to i is sufficient for a thing to have P .

¹ Although, as we explain in chapter 8 above, the proposition expressed by the claim in the first situation remains true. The sentence does not express the same proposition in the second situation.

Dunn suggests that ‘*i*’s having *P* essentially’ should be parsed as ‘*i*’s having *P* both relevantly and necessarily’ (Dunn 1990). Dunn does this in order to avoid problems with the standard treatment of essential predication as mere necessary predication. Suppose that we follow the standard path and say that *i* has *P* essentially if and only if *i* has *P* in all possible worlds in which *i* exists. Then suppose that *A* is some necessary truth. For any object *i*, *i* has the predicate $\lambda x.A$ essentially. Thus, everything is such that either Ramsey is furry, Ramsey is not furry, or Ramsey does not exist.

Dunn’s parsing of essential predication in terms of necessary relevant predication, however, is not unambiguous. He gives us three different formalisations of it (ibid., p 83), but the one that I think the best candidate for essential predication is

$$\Box \forall x (x = a \rightarrow P(x)).$$

This says that in any accessible situation, *x* has *P* relevantly. It would seem reasonable to treat the necessity here as metaphysical necessity.

Dunn goes on to discuss essential relational properties. Consider a set, $X = \{Ramsey\}$. Being in this set is not essential to Ramsey, but having Ramsey as a member is essential to the set. For the identity of a set is determined by its members. So, we have

$$\Box \forall x (x = X \rightarrow ramsey \in x),$$

but not

$$\Box \forall x (x = ramsey \rightarrow x \in X).$$

In the standard possible world theory, this distinction cannot be made.

11.4 Relevance and belief revision

In (Mares 2002a), I set out a theory of belief revision that uses relevant logic. I won’t present the theory in any detail here, for it is technically quite complicated. Interested readers should consult the paper for the full working out of the theory. But, in broad outline, the idea is quite simple.

As we have seen, relevant logic is a form of paraconsistent logic. Once we adopt a paraconsistent logic, we open the door to the possibility that an agent can rationally and knowingly believe a contradiction.

But there is a problem here. Traditional theories of belief revision are largely theories about how to maintain consistency in the light of new data, that perhaps contradicts our beliefs. If we allow contradictory beliefs in our agents, should our theory of belief revision degrade into mere data collection?

Here is an example. Suppose I am in the house. My friend Jane is visiting me. I believe that Ramsey is out in the garden. I ask Jane to call him to get

him inside. She tells me that he is not in the garden. Should I now believe that Ramsey is both in the garden and that he is not in the garden? Clearly, not. I should reject the old belief that he is in the garden and maintain (at least as far as Ramsey's whereabouts are concerned) consistent beliefs.

Now we have a problem. We need to give a theory of reasons for relinquishing beliefs. To this end, I use the theory of content that we discussed in chapter 9 above. The process can be simply described as follows. Suppose that an agent is incorporating a new proposition into her belief set. If the resulting belief set is pragmatically consistent with her rejection set, then there is no problem. If it is not pragmatically consistent, then she will either have to down-size her old stock of beliefs or denials or both.

Let's say that our agent is trying to incorporate a new statement, say, A , into her belief set and that it creates a quasi-inconsistency. Then there will be a set of pairs (B, D) such that B is from her belief set, D is from her denial set, and it is a theorem of our logic that $(A \wedge B) \rightarrow D$. In order to remove the quasi-inconsistency, the agent will have to remove at least one member of each of these pairs. She should do so on the basis of which deletions will provide her with the best overall resulting structure of beliefs and denials. To make this choice our ideal agent would use the usual list of theoretical virtues, such as, elegance, simplicity, conservativeness (i.e., be reluctant to delete beliefs or rejections), and explanatory coherence.

Also, as I have argued in (Mares 2000) and in chapter 9 above, I think that we should reject contradictions by default. We have good inductive reasons to accept this default. Rejecting contradictions by and large has resulted in the past from the adoption of correct beliefs. Given this default, we can understand why I should reject my belief that Ramsey is outside when I hear from Jane that he is not outside. By default I reject the statement that Ramsey is both outside and not outside. When I add the belief that Ramsey is not outside, my beliefs and rejections are quasi-inconsistent. Let ' O ' stand for 'Ramsey is outside.' Then I have a rejection of $O \wedge \sim O$. I also have a prior belief that O and am incorporating a new belief that $\sim O$. Since it is a theorem that $(O \wedge \sim O) \rightarrow (O \wedge \sim O)$, we have a pragmatic inconsistency, that is, my belief set implies something in my rejection set. The obvious choice is to abandon the belief in $\sim O$, since there seems to be no reason to override the default rejection of the contradiction.

What is new about this theory can be seen when a contrast is made to traditional theories of belief revision. On the popular theory developed by Carlos Alchourón, Peter Gärdenfors and David Makinson (AGM), the two driving mechanisms of belief revision are conservativeness and consistency. Here we consider only the version of AGM that uses 'epistemic entrenchment'. According to AGM when we incorporate a new belief that is inconsistent with our old beliefs, we reject those old beliefs that are less entrenched until we achieve a

new belief set that is consistent. When we get rid of consistency, what do we replace it with? My answer is to replace it with pragmatic consistency.

11.5 Reasoning about fiction

One way in which the foregoing theory of belief revision can be made to do some work is in modelling our reasoning about fictional stories. The theory I outline in this section is rather sketchy and is not a complete theory, but rather is a work in progress.

When we reason about a story, we apply background theories. For example, consider the following, rather unremarkable, passage from a novel:

I had to get some air. I locked the door behind me and left the motel. I crossed the street and sat down on the sea wall, staring down at the stretch of beach where Jean Timberlake had died. Behind me, Floral Beach was laid out in miniature, six streets long, three streets wide. ((Grafton 1989), p. 211)

The author is trying to set a mood here by giving a description. In our imaginations, we fill out her picture. We think of streets with shops or houses, say, perhaps some surf rolling in, and so on. Suppose we are asked questions about the passage, such as, 'Is the sea filled with salt water or formaldehyde?' We would answer 'salt water'. And 'What colour are the roads?' We would answer 'black', not 'neon pink'. And so on. In understanding a story we often fill in details based on our understanding of our own world. But not all of these details come from our theories about what is actually true. Consider, for example, science fiction stories in which space craft can accelerate beyond the speed of light. To understand these stories we often fill details with our understanding of the 'space-time warp' or 'hyperspace' gained from reading other science fiction stories.

The picture that I have been sketching is drawn from the philosophy of science. When scientists do an experiment to test a theory, they apply both the theory that is under test and background theories. When we read² further into a story, we apply both our understanding of the events described by the story from what we have read so far and our background theories. When they find data that does not cohere with their collection of theories (the theory being tested or the current view of the events of the story and the background theories), scientists and readers alike are stuck with the same problem. Should they reject some elements of their current view of the events described or the theory under test, or should they reject a background theory? In the philosophy of science,

² Nothing in this section crucially depends on a story's being read rather than its being seen presented in a film or a play or heard in a radio play, and so on.

this is known as the ‘Duhem–Quine problem’, named after Pierre Duhem and W. V. O. Quine. The Duhem–Quine problem as it is applied to fiction, however, is in some ways more difficult than the problem for scientists. For it is sometimes the case in fiction that we can accept contradictory propositions, but it is often the case that we would rather not do so.

This is where my theory of belief revision can be of some use. Let us look at two examples to see how we can use the theory. The first example is from the cartoon series *South Park*. In almost every episode of the early seasons of the series, the same character is killed. The series, however, is a continuing story. The character who dies – Kenny – is alive again at the beginning of the next episode. This happens largely without comment, but sometimes the other characters notice that this happens and once declare that this does not make sense. Our background belief that dead people do not come back to life is used to interpret other parts of the stories. In my opinion, what we do is accept the contradiction. We recognise that this contradiction is intended by the creators of the series and we take it on board as part of the comic content of the programme.

On the other hand, consider the story ‘Diary of a Madman’ by Nicolai Gogol. In it, the main character and narrator reports hearing dogs discussing him. I discount this information rather quickly, since we are told by the title of the story that we are dealing with a madman, and I take the events described to be the figment of his imagination. We do so despite the fact that Gogol is an absurdist writer and is not adverse to writing inconsistent stories. Of course we do construct the content of the inner life of the main character as well in reading the story and this inner life constitutes a fiction within a fiction. For the construction of the character’s inner life, we no longer include the proposition that dogs cannot talk, since this construction requires that we adopt the background beliefs and rejections of the character himself.

11.6 Relevant logic and the foundations of mathematics

We saw in chapter 9 that we can combine relevant logic with a theory of reasonable inference to vindicate classical reasoning. But we can also formulate theories in relevant logic without the additional trappings of denial, and so on. The resultant theories have been the focus of some very interesting mathematics and philosophy. There are various ways in which these theories are philosophically interesting. For example, it is interesting to see how much of mathematical reasoning can be treated as relevant inference. It is also interesting to see what is lost if we get rid of certain classical principles of reasoning. In the next few sections we will look at relevant mathematical theories and their properties to see what philosophical conclusions, if any, can be drawn from them.

11.7 Relevant arithmetic

One especially interesting relevant theory is Bob Meyer's $R^\#$, a relevant form of Peano arithmetic. Meyer's aim was to show that we can reconstruct arithmetical reasoning relevantly. As we shall see below, relevant Peano arithmetic, or $R^\#$, has some very interesting properties. Unfortunately, $R^\#$ didn't have all the properties that Meyer had hoped it had.

In order to reconstruct arithmetical reasoning, Meyer wanted to show that all of the theorems of classical Peano arithmetic could be proven in $R^\#$. He proved various results, some with very elegant proofs, to this end. In particular, he showed that if we add the gamma rule to $R^\#$ it contains all of classical Peano arithmetic (see (Friedman and Meyer 1992)). The gamma rule, as we remember, is a version of disjunctive syllogism, namely,

$$\frac{\begin{array}{c} \vdash A \vee B \\ \vdash \sim A \end{array}}{\vdash B}.$$

From the theoremhood of $A \vee B$ and $\sim A$ we can infer that B is also a theorem.

Meyer also showed that $R^\#$ enriched by the addition of the omega rule contains all of classical Peano arithmetic (also with the addition of the omega rule). The omega rule is an infinitary rule of inference that states that if a property P is such that $P(0)$, $P(1)$, $P(2)$, and so on, we can infer that $\forall x P(x)$. Meyer called the system that results by adding the omega rule to $R^\#$, $R^{\#\#}$. He was able to show, by means of a lovely argument, that the gamma rule is admissible in $R^{\#\#}$ and following from this that all of classical Peano arithmetic is contained in $R^{\#\#}$. The fact that $R^{\#\#}$ contains an infinitary rule of inference, however, is a flaw. We cannot use infinitary rules. We can never state all the premises.

$R^\#$ itself did not admit gamma. Through a complicated series of proofs, Meyer and Harvey Friedman showed that gamma is not admissible in $R^\#$ and, following from this, that $R^\#$ does not contain all of classical Peano arithmetic. More seriously, the Friedman–Meyer result indicates that there are important parts of number theory not included in $R^\#$ (Friedman and Meyer 1992).

A current project in relevant arithmetic, therefore, is to find a system strong enough to contain all of classical Peano arithmetic but without the use of suspect devices such as infinitary rules of inference. This as yet fictional system has been dubbed ' $R^{\#1/2}$ ' by Meyer.

11.8 Relevant arithmetic and Gödel's second theorem

Despite the fact that $R^\#$ is inadequate to capture all of standard arithmetical reasoning, it is a very interesting theory. What is perhaps most interesting

about it is that the ‘absolute consistency’ of $R^\#$ can be proved using elementary methods.

To explain what this means and its importance, we should briefly discuss Hilbert’s programme in the foundations of mathematics. Hilbert wanted to provide a strict finitist foundation for mathematics. On Hilbert’s view, the basis for our mathematics is our finitary combinatorial intuitions. These intuitions tell us only about finite numbers, finite collections, finitary operations and rules, and so on. We can talk about infinite sets and use them in our mathematics, but what is important about them is what they allow us to prove about finite mathematics. Thus, our infinitary mathematics has to be a *conservative extension* of our finitary mathematics in order to be considered safe.

Georg Kreisel has put this neatly in the following way. Hilbert’s problem is to show that, for a given system of mathematics,

$$\text{Pr}(\ulcorner A \urcorner) \rightarrow A$$

where A is a statement about finitary mathematics and ‘Pr’ is a provability predicate for the system in question ((Kreisel 1967), p. 233).

Gödel’s second theorem states that certain systems of arithmetic cannot prove themselves consistent. Thus, it is impossible to show in these systems that $\sim\text{Pr}(\ulcorner f \urcorner)$, where f is a falsum for classical logic. This formula is equivalent to

$$\text{Pr}(\ulcorner f \urcorner) \rightarrow f,$$

where the arrow is material implication. Now, in Peano arithmetic and related systems, the formula $0 = 1$ is equivalent to f . So we cannot show that

$$\text{Pr}(\ulcorner 0 = 1 \urcorner) \rightarrow 0 = 1.$$

But ‘ $0 = 1$ ’ is a statement of finite arithmetic. Hence, it would seem that Hilbert’s programme fails for Peano arithmetic.

When we move from Peano arithmetic to $R^\#$ something interesting happens. The property that a system has if it cannot be proven in it that $0 = 1$ is called ‘absolute consistency’. Meyer has shown using finitary methods that $R^\#$ is absolutely consistent (Meyer 1976), (Meyer and Mortensen 1984). Thus, $R^\#$ has a definite advantage over Peano arithmetic. But what is perhaps even more interesting is trying to figure out what it is about $R^\#$ that allows us to prove its absolute consistency using finitary means.

Note that Meyer has not proven that Gödel’s second theorem fails in relevant arithmetic. Although he has used finitary means to prove his result, he has not shown that this proof can be proven in $R^\#$ itself. Moreover, it is not clear whether there is a provability predicate in $R^\#$ that has similar properties to the provability predicates of classically based systems of arithmetic. So, it is an open question about the status of Gödel’s second theorem in relevant arithmetic. And, perhaps

even more interestingly, there is the question whether we can base a partial realisation of Hilbert's programme on $R^\#$.

11.9 Relevant set theory?

Producing a relevant version of arithmetic seems like a worthwhile project. Arithmetic is an important human activity and, if we are to have a theory of reasoning, it would be nice to capture that sort of reasoning that goes on when we do arithmetic.

From the late nineteenth century, set theory has gained prominence as a foundation for mathematics. Thus, it would also seem reasonable to want to reconstruct set theory so that it is based on relevant rather than classical logic. Some relevant logicians have tried to do just that, most notably Ross Brady. But some other relevant logicians see this attempt as a mistake. In particular, Bob Meyer has held³ that set theory is intrinsically wedded to classical logic. Set theory grew up with classical logic in the late nineteenth and early twentieth centuries. It was meant to capture the notion of an extension. Set theory, thus, should be connected to an extensional logic like classical logic, and not an intensional one like relevant logic.

Consider the usual definition of the subset relation:

$$X \subseteq Y =_{df} \forall x(x \in X \supset x \in Y).$$

How should we define subset in a relevant set theory? We can't use the material conditional. It is too weak. For suppose that we did use the above definition, with the material conditional (defined as $A \supset B =_{df} \sim A \vee B$). Then we could have a situation s in which $X \subseteq Y$ and an individual a in X but fail to be in Y , since modus ponens for material implication does not hold in all situations. Surely the subset relation is stronger than that.

The obvious alternative is to change the definition to

$$X \subseteq Y =_{df} \forall x(x \in X \rightarrow x \in Y).$$

This definition, on the other hand, seems just too strong. Consider the statement

Every woman I have met named 'Ismay' lives in Seatoun.

This statement is true. I have met two women named 'Ismay' and they both live in the Wellington suburb of Seatoun. Let I be the set of women named 'Ismay', M be the set of people that I have met, and S be the set of people living in Seatoun. If we use the above definition of subset, we would be committed to holding that

$$\forall x((x \in I \wedge x \in M) \rightarrow x \in S).$$

³ In seminars and in personal communication.

This seems extremely strong. Recall our interpretation of relevant implication. For a relevant implication to be true, there must be some real informational link between the antecedent and the consequent (or some logical connection derived from the rules of the natural deduction system). But it just happens that I have met only two women with that name and they both live in Seatoun.

One might think that there is a logical link between a subset and set that is strong enough to support an implication. But this is not the case. Assume that we have in our language a set abstractor, λ . If A is a sentence in first-order relevant logic, perhaps with the variable x free, then ' $\lambda x A$ ' denotes the set of entities which satisfy A . Let p and q be propositional variables. At any worldly situation,

$$\lambda y p \subseteq \lambda y (q \vee \sim q).$$

Thus, by our tentative definition of subset, it would seem that the following is also true:

$$\forall x (x \in \lambda y p \rightarrow x \in \lambda y (q \vee \sim q)).$$

Thus, by universal instantiation (where a is some individual in the worldly situation),

$$a \in \lambda y p \rightarrow a \in \lambda y (q \vee \sim q).$$

By lambda conversion, we then have,

$$p \rightarrow (q \vee \sim q),$$

which is a paradox of strict implication and a thesis we want to reject.⁴

Despite this problem, I think that a relevant set theory is both needed and possible. Consider again the sentence 'Every woman I have met named "Ismay" lives in Seatoun'. How should we formulate it (not its set-theoretic translation) in relevant logic? Not as either of,

$$\forall x ((M(x) \wedge I(x)) \rightarrow S(x))$$

or,

$$\forall x ((M(x) \wedge I(x)) \supset S(x))$$

for the reasons given above. It is clear that any logic should have a means for formalising categorical propositions.

⁴ An obvious way around this argument is to reject unrestricted lambda abstraction and conversion. For example, we could use Dunn's theory of relevant predication, discussed above, to restrict our theory to sets determined by relevant predicates. But then very difficult questions would arise, such as whether we could include the empty set in our theory.

Whether an adequate foundations of mathematics can be produced in this way, I do not know.

One means for doing so is due to Nuel Belnap and Dan Cohen. Cohen has added Belnap's conditional assertion connective to relevant logic ((Anderson *et al.* 1992), §75). The expression,

$$B/A,$$

means that A is asserted on the condition that B . Or, if B is true, then A is asserted. If B is false (or, more accurately, not true)⁵ then A is not asserted. Belnap and Cohen have argued that 'All X 's are Y 's' should be read ' $\forall x(X(x)/Y(x))$ '. For everything, if it is an X , then we assert of it that it is also a Y . If it is not an X , then we need not assert anything of it.

With the addition of conditional assertion to relevant logic, we can formalise categorical statements and, I think, statements about subsets. I suggest that,

$$X \subseteq Y =_{df} \forall x(x \in X / x \in Y)$$

is a reasonable definition of subset. What axioms would be needed for an adequate set theory of this sort, I do not know just yet. But I think it is a promising direction for research.

11.10 Weaker systems of relevant logic

In this book, we have examined the relevant logic R. This is one of the strongest relevant logics. One logic is said to be 'stronger' than another if and only if the first includes all the theorems of the second and other theorems as well. In the final sections, we will examine weaker systems of relevant logic.

There are a great many weaker systems of relevant logic that have been studied and motivated. One of the most interesting is the logic S (for 'syllogism'). This logic has two axioms, viz.,

$$(B \rightarrow C) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$$

and

$$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$$

and the rule of modus ponens. (The only connective in the language of this system is implication.) As Errol Martin has shown, this system has no theorems of the form $A \rightarrow A$. In terms of inference, according to the system S, all inferences of a proposition from itself are invalid (Martin and Meyer 1982). Thus,

⁵ The semantics for conditional assertion is three valued. We will ignore this here, but I think that the form of three valuedness required can be incorporated into our framework without any real damage to our view.

Martin, together with Meyer, maintains that it is a virtue of *S* that it rejects the traditional fallacy of circular reasoning.⁶

Another very weak relevant logic is the logic *B*. In contrast to *S*, the theorems of *B* that contain implication as their only connective are *all* of the form $A \rightarrow A$.

Most relevant logicians accept logics stronger than *B*. In the next section, we will discuss the interpretation and motivation of some of these systems.

11.11 Brady's content semantics

One way of understanding some weaker systems of relevant logic is through Ross Brady's content semantics. Brady's full formal theory is very complicated, but we can get the flavour of it rather quickly here. The view is set out in full in (Brady 2003), §12.5 and in (Brady forthcoming).

We start with the sentences of some language. This is not a formal language, but rather what we will call an 'interpretational language'. It needs to be a language like English or the language of standard mathematics, i.e., a language that is already interpreted.⁷ Suppose that X is a set of sentences of the interpretational language. Then $c(X)$ is the content of that sentence. The content consists of all of the sentences that can be 'analytically established' from the members of X ((Brady 2003), p. 263). Say, for example, that X is just the singleton {John is a bachelor}. Then the sentence 'John is unmarried' is in the content $c(X)$. When X is a singleton $\{S\}$, we write $c(S)$ instead of $c(\{S\})$.

The meanings of the various connectives can be given in terms of contents. Disjunction is the most straightforward. The content of a disjunction $c(S \text{ or } T)$ is $c(S) \cap c(T)$. This might seem wrong. For usually disjunction is paired with set theoretic union rather than intersection. But intuitively, for contents, the pairing of disjunction and intersection is right. Think of the content of the disjunction 'Table x is made of rimu or table x is 12 feet wide.' In the content of 'Table x is made of rimu' is the sentence 'This table is made of wood' but not 'Table x is more than 11 feet wide.'² In the content of 'Table x is 12 feet wide' is the sentence 'Table x is more than 11 feet wide' but not the sentence 'Table x is made of wood.' We cannot establish either 'Table x is made of wood' or 'Table x is more than 11 feet wide' from the original disjunction. Thus, neither belongs in the content of the disjunction, although their disjunction ('Table x is made of wood or table x is more than 11 feet wide') does belong in that content.

The content of a conjunction $c(S \text{ and } T)$ is $c(c(S) \cup c(T))$, i.e., the content of the unions of the contents of each conjunct. We can make sense of the content

⁶ The philosophical motivations for *S*, including those presented here are in an as yet unpublished manuscript by Martin and Meyer called 'S is for Syllogism'.

⁷ Brady does not use the phrase 'interpretational language'. This is my invention. In fact, although Brady motivates his semantics using an interpretational language, his model theory is very algebraic in nature, taking contents to be primitive elements of the structure.

of conjunctions as closures of unions using a similar argument to the one that established the connection between disjunction and intersection.

The content of an implication is more interesting. In our interpretational language, we include both names of the various contents and the set theoretic relation $\supseteq$. The relation $\supseteq$ is the superset relation, that is, if X and Y are sets, then $X \supseteq Y$ if and only if $Y \subseteq X$. There are various important analytic inferences that govern the behaviour of the superset relation. In particular, consider the set $\{c_1 \supseteq c_2, c_2 \supseteq c_3\}$. In the content of this set $c(\{c_1 \supseteq c_2, c_2 \supseteq c_3\})$ there is the sentence $c_1 \supseteq c_3$, since superset is a transitive relation. We shall see why this fact is important soon. The content of an implication ' S implies T ' is the set $c(c(S) \supseteq c(T))$. This content includes (and only includes) what can be analytically inferred from the sentence ' $c(S) \supseteq c(T)$ '.

Now we bring in a formal language, that includes propositional variables, conjunction, disjunction, implication and parentheses. We should also include negation, but it introduces a host of difficulties that we wish to avoid here. An interpretation I assigns to each formula of our language a content. Given an assignment of contents to propositional variables, the clauses for the connectives are the following:

- $I(A \vee B) = I(A) \cup I(B)$;
- $I(A \wedge B) = c(I(A) \cup I(B))$;
- $I(A \rightarrow B) = c(I(A) \supseteq I(B))$.

We can see that the logic that this semantics characterises is quite a bit weaker than our favourite system, R. As we have seen, implication is transitive in Brady's logic, which is called DJ^d , but in a weaker sense than R. In R, we get as a theorem,

$$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C)).$$

But the transitivity that is captured by the content semantics only justifies the scheme,

$$((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C).$$

More importantly, as we saw in chapters 2 and 3, our semantics makes valid the scheme of contraction, i.e.,

$$(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B).$$

The contraction scheme, however, cannot be justified using the content semantics. For to do so would be to show that $c(c_1 \supseteq c(c_1 \supseteq c_2)) \supseteq c(c_1 \supseteq c_2)$, for all contents c_1 and c_2 , and this is not possible. As we shall see soon, the failure of contraction is a key feature of Brady's logic.

As we can see, Brady's semantics provides an intuitive interpretation of a logic. As in the case with Read's interpretation of relevant logic, it does not seem appropriate for me to criticise it here. It is a very different approach from

mine. Moreover it is not clear to me that we cannot accept Brady's semantics as well as my own. We merely take his as defining a very different notion of implication.

There are other interpretations of weak relevant logic. For example, as we have already seen, there is Priest's use of 'logic fictions' to interpret a weak relevant logic (see §3.10 and see (Priest 2001)). His model consists in a real world and a set of non-normal worlds. An implication, ' $A \rightarrow B$ ', is true in the real world if and only if for all non-normal worlds w , if ' A ' is true in w then ' B ' is also true in w . For non-normal worlds, implications have no set truth conditions – they are made true or false randomly. These non-normal worlds are like the universes of science fiction, except that they reject laws of logic rather than laws of physics. The logic characterised by this semantics is much weaker than Brady's systems. I criticised Priest's semantics briefly, but it does have the virtues of being very simple and elegant.

11.12 Relevance and naive theories

Brady and others have used weaker relevant logics, like DJ^d , to formulate naive set theories, naive theories of truth, and naive higher-order logics. Naive set theories and naive higher-order logics contain naive comprehension principles.

A naive comprehension principle for set theory is a principle like the following:

$$\exists x \forall y (A(y) \leftrightarrow y \in x),$$

where y is the only variable free in $A(y)$. The naive comprehension principle says that for any open formula $A(y)$ there is a set x the members of which are exactly those things that satisfy the open formula.

The naive comprehension principle is very intuitive. It tells us that every formula of set theory determines a set. But adding it to standard set theories, which are based on classical logic and even relevant logics like our favourite system R , has produced inconsistencies. For example, Russell's paradox can be derived in naive set theory (i.e., a set theory with naive comprehension) based on R as follows. Suppose that we have a language that includes quantifiers (which range over the domain of sets) and the relation of set membership ($\in$). First, we replace $A(y)$ in the comprehension principle with $\sim y \in y$. Now we can produce the following derivation:

- | | |
|---|---------------------|
| 1. $\exists x \forall y (\sim y \in y \leftrightarrow y \in x)_{\emptyset}$ | naive comprehension |
| 2. $\forall y (\sim y \in y \leftrightarrow y \in r)_{\emptyset}$ | 1, $\exists E$ |
| 3. $\sim r \in r \leftrightarrow r \in r_{\emptyset}$ | 2, $\forall E$ |

The reason we use the empty set as a subscript throughout is that our assumption is not a normal hypothesis. Rather, we are assuming that naive comprehension

is an axiom of our logic; if it were an axiom, we would be able to put it into any proof with the empty set as its subscript. In the derivation we prove the existence of a set that belongs to itself if and only if it does not belong to itself.

Note that we do not yet have a contradiction. We have a paradoxical biconditional. We can turn this paradox into an out-and-out contradiction if we add certain other principles of logic, such as the principle of excluded middle, which in this case tells us that either r is or is not a member of itself. Thus, in logics with excluded middle, such as our favourite logic R , we obtain a contradiction. Weaker systems, like Brady's logic, which reject excluded middle and other related principles, can act as the basis for a consistent naive set theory (see (Brady forthcoming)).

There are, however, philosophers who accept an inconsistent naive set theory. These philosophers come in two camps. There are those like Chris Mortensen who adopt inconsistent set theory because they think that it is useful. That is, they are instrumentalists about inconsistent set theory (Mortensen 1995). On the other hand, there are realists like Graham Priest who hold that there is a universe of sets and that an inconsistent set theory accurately represents that universe ((Priest 1987), chapter 2).

Whichever approach we take to naive set theory, we have to beware of worse possible consequences than a few contradictions. For if we add the naive comprehension principle to a set theory based on the logic R , we obtain a system that is *trivial*. That is, we can derive not just a contradiction, but any formula whatsoever in this theory. Let p be an arbitrary formula. The following proof shows that we can derive p in a naive set theory based on R .

1. $\exists x \forall y ((y \in y \rightarrow p) \leftrightarrow y \in x)_{\emptyset}$	naive comprehension axiom
2. $\forall y ((y \in y \rightarrow p) \leftrightarrow y \in c)_{\emptyset}$	1, $\exists E$
3. $(c \in c \rightarrow p) \leftrightarrow c \in c_{\emptyset}$	2, $\forall E$
4. $c \in c \rightarrow (c \in c \rightarrow p)_{\emptyset}$	3, $\wedge E$
5. $(c \in c \rightarrow (c \in c \rightarrow p)) \rightarrow (c \in c \rightarrow p)_{\emptyset}$	contraction
6. $c \in c \rightarrow p_{\emptyset}$	4, 5, $\rightarrow E$
7. $(c \in c \rightarrow p) \rightarrow c \in c_{\emptyset}$	3, $\wedge E$
8. $c \in c_{\emptyset}$	6, 7, $\rightarrow E$
9. $p_{\emptyset}$	6, 8, $\rightarrow E$

In this manner, we can prove any formula whatever. The key step in this proof is the use of the principle of contraction in step 5. The principle of contraction, in schematic form is

$$(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B).$$

As we saw in chapters 2 and 3 above, I accept contraction. Here I will briefly repeat my defence of this acceptance.

The contraction principle corresponds closely with the ability to use a hypothesis more than once in a proof. Here is a proof of contraction that shows this correlation:

1. $A \rightarrow (A \rightarrow B)_{\{1\}}$ *hyp*
2. $A_{\{2\}}$ *hyp*
3. $A \rightarrow B_{\{1,2\}}$ $1, 2, \rightarrow E$
4. $B_{\{1,2\}}$ $2, 3, \rightarrow E$
5. $A \rightarrow B_{\{1\}}$ $2 - 4, \rightarrow I$
6. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)_{\emptyset}$ $1 - 5, \rightarrow I$

What allows the derivation of contraction is the ability to use the second hypothesis twice in implication eliminations, both in the derivation of line 3 and in the derivation of line 4. One way of barring contraction is to treat the subscripts as multisets (as in Stephen Read's Scottish Plan discussed in chapter 9 above). If we do so, the closest that we can get is the following proof:

1. $A \rightarrow (A \rightarrow B)_{[1]}$ *hyp*
2. $A_{[2]}$ *hyp*
3. $A \rightarrow B_{[1,2]}$ $1, 2, \rightarrow E$
4. $B_{[1,2,2]}$ $2, 3, \rightarrow E$
5. $A \rightarrow B_{[1,2]}$ $2 - 4, \rightarrow I$
6. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)_{[2]}$ $1 - 5, \rightarrow I$

The square brackets indicate that we are dealing with multisets rather than sets. The derivation does not prove contraction to be a theorem. Rather, it shows that if we assume the antecedent of contraction is true at one situation, we can show that there is some situation at which the whole contraction thesis is true. To eliminate the reference to situation 2 in the conclusion we can hypothesise the formula A a second time, as in the following argument:

1. $A \rightarrow (A \rightarrow B)_{[1]}$ *hyp*
2. $A_{[2]}$ *hyp*
3. $A \rightarrow B_{[1,2]}$ $1, 2, \rightarrow E$
4. $A_{[3]}$ *hyp*
5. $B_{[1,2,3]}$ $3, 4, \rightarrow E$
6. $A \rightarrow B_{[1,2]}$ $4 - 5, \rightarrow I$
7. $A \rightarrow (A \rightarrow B)_{[1]}$ $2 - 6, \rightarrow I$
8. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow (A \rightarrow B))_{\square}$ $1 - 5, \rightarrow I$

Here we have the empty multiset ($\square$) as a subscript in the conclusion, but we do not have the formula that we want. What we get is just a complex form of the identity axiom ($A \rightarrow A$).

What happens in these proofs is that we must keep track not only of the premises that we use but how many times we use them. There is something

mathematically very elegant about this approach to natural deduction. Brady's content semantics, presented in the previous section, gives an interpretation of a contraction-free natural deduction system (Brady forthcoming), and John Slaney (Slaney 1990) and Stephen Read (Read 1988) give interpretations of other systems that reject contraction. Whether any of their justifications show that there is something illegitimate about using premises as often as we wish in ordinary inference I have my doubts. But I will not argue that point here. I only wish to point the reader to this very interesting literature. The Slaney and Read systems, however, are not weak enough to support a non-trivial naive set theory. I only discuss them here to point to the use of methods we have seen earlier in the book to reject contraction.

11.13 Summary of part III

In part III of the book, we tidied up some loose ends. In chapter 9, we covered various technical topics to do with deduction. This set up the discussion in chapter 10 concerning what is perhaps one of the most difficult issues regarding relevant logic – its rejection of the rule of disjunctive syllogism. I argued in chapter 10 that, in fact, we can accept disjunctive syllogism in certain conditions and that this restriction on the acceptability of disjunctive syllogism is not untoward. In chapter 11, I end the book by discussing uses of relevant logic including some uses of systems weaker than the logic R, the system that has been the main focus of the rest of the book.

This book is an interpretation and defence of relevant logic. These two projects of interpreting and defending relevant logic are not distinct. For it has been the chief complaint about relevant logic that it has no reasonable interpretation.

My interpretation has used as its central notion the real use of premises in a deductive inference. This, I think, makes sense. For it makes *relevance* the key notion in understanding relevant logic. We saw this notion at work in the understanding of relevant deduction, situated inference, the theory of implication, and, in a somewhat weaker sense, in the theory of conditionals. In other words, relevance as the real use of premises (or the real use of antecedents, in the case of implications and conditionals) permeates my whole interpretation of relevant logic.

I also have tried to short-circuit metaphysical criticisms of relevant logic by constructing a model for the logic from an ontology that is very close to that standardly used by modal logicians. The one difference is that my ontology uses non-well-founded set theory rather than standard set theory. But this is a minor difference, for the theory can be reconstructed using standard set theory, albeit in a less elegant fashion, and, besides, there are other very good reasons for accepting non-well-founded set theory. The upshot of all of this is that we can use relevant logic without any pangs of philosophical guilt. It has a reasonable interpretation and it does not commit us to an extravagant ontology any more than modal logic does.

Nor should we feel guilty about the alleged logical weakness of relevant logic. I have argued in chapter 10 that there are various ways that we can interpret disjunctive syllogism – which we think is invalid in the strict formal sense – to make it into a reasonable form of inference from a relevant point of view.

I end the book by discussing uses of relevant logic, as a basis for deontic logic, as a basis for a theory of belief revision, and as a part of the foundations of mathematics. These programmes are far from complete and so we end both by saying that relevant logic is useful and that there is plenty of work for logicians to do on relevant logic and its applications.

Appendix A The logic R

A.1 THE PROPOSITIONAL LANGUAGE

The propositional vocabulary consists of

- Propositional variables: $p, q, r, \dots$
- A unitary connective: $\sim$
- Two binary connectives: $\wedge, \rightarrow$
- Parentheses: $(,)$.

We use the usual formation rules.

A.2 DEFINED CONNECTIVES

- $A \vee B =_{df} \sim(\sim A \wedge \sim B)$
- $A \leftrightarrow B =_{df} ((A \rightarrow B) \wedge (B \rightarrow A))$

A.3 AXIOMATISATION

A.3.1 AXIOM SCHEMES

1. $A \rightarrow A$
2. $(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$
3. $A \rightarrow ((A \rightarrow B) \rightarrow B)$
4. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$
5. $(A \wedge B) \rightarrow A$
6. $(A \wedge B) \rightarrow B$
7. $A \rightarrow (A \vee B)$
8. $B \rightarrow (A \vee B)$
9. $((A \rightarrow B) \wedge (A \rightarrow C)) \rightarrow (A \rightarrow (B \wedge C))$
10. $(A \wedge (B \vee C)) \rightarrow ((A \wedge B) \vee (A \wedge C))$
11. $\sim\sim A \rightarrow A$
12. $(A \rightarrow \sim B) \rightarrow (B \rightarrow \sim A)$

A.3.2 RULES

$$\begin{array}{c}
 \vdash A \\
 \vdash B \\
 \hline
 \therefore \vdash A \wedge B \quad \wedge I \\
 \vdash A \rightarrow B \\
 \vdash A \\
 \hline
 \therefore \vdash B \quad MP
 \end{array}$$

A.4 NATURAL DEDUCTION SYSTEM

These rules are taken, with slight modification, from (Anderson *et al.* 1992), p. xxvi.

A.4.1 STRUCTURAL RULES

- A formula may be introduced as a hypothesis of a new subproof. Each new hypothesis receives a subscript of the form $\{k\}$ where k has not been used previously in the proof. (hyp)
- A_α may be repeated. (rep)
- A_α may be repeated into a hypothetical subproof. (reit)

A.4.2 IMPLICATION RULES

- From a proof of B_α on the hypothesis that $A_{\{k\}}$ to infer that $A \rightarrow B_{\alpha-k}$, provided that k is in α . ($\rightarrow I$)
- From $A \rightarrow B_\alpha$ and A_β to infer $B_{\alpha \cup \beta}$. ($\rightarrow E$)

A.4.3 CONJUNCTION RULES

- From A_α and B_α to infer $A \wedge B_\alpha$. ($\wedge I$)
- From $A \wedge B_\alpha$ to infer A_α and from $A \wedge B_\alpha$ to infer B_α . ($\wedge E$)

A.4.4 DISJUNCTION RULES

- From A_α to infer $A \vee B_\alpha$ and from B_α to infer $A \vee B_\alpha$. ($\vee I$)
- From $A \vee B_\alpha$, $A \rightarrow C_\beta$, and $B \rightarrow C_\beta$, to infer $C_{\alpha \cup \beta}$. ($\vee E$)

A.4.5 DISTRIBUTION RULE

From $A \wedge (B \vee C)_\alpha$, infer $(A \wedge B) \vee (A \wedge C)_\alpha$.

A.4.6 NEGATION RULES

- From $A \rightarrow f_\alpha$ to infer $\sim A_\alpha$. ($\sim I$)
- From $\sim A_\alpha$ to infer $A \rightarrow f_\alpha$. ($\sim E$)
- From $\sim \sim A_\alpha$ to infer A_α . (DN)

Appendix B Routley–Meyer semantics for R

Routley and Meyer first presented their semantics in (Routley and Meyer 1973). This presentation takes some elements from (Anderson *et al.* 1992) and some from Dunn’s semantics for negation (Dunn 1993).

B.1 FRAMES

An R-frame is a quadruple $\langle Sit, Logical, R, C \rangle$ such that Sit is a non-empty set, $Logical$ is a non-empty subset of Sit , and R is a three-place arbitrary place relation on Sit , C is a binary relation on Sit , which satisfies the following definitions and postulates:

- $s \leq t =_{df} \exists x(x \in Logical \ \& \ Rxst)$;
- $R^1stu =_{df} Rstu$;
- $R^{n+1}s_1 \dots s_{n+2}t$ iff $\exists x(R^n s_1 \dots s_{n+1}x \ \& \ Rxs_{n+2}t)$, for $n \geq 1$;
- if $R^n \dots s_i s_{i+1} \dots t$, then $R^n \dots s_{i+1} s_i \dots t$;
- if $R^n \dots s_i \dots t$, then $R^{n+1} \dots s_i s_i \dots t$;
- $Rsss$;
- for each situation s there is a unique $\leq$ -maximal situation s^* such that Css^* ;
- if $Rstu$, then Rsu^*t^* ;
- $s^{**} = s$;
- if $s \leq t$, then $t^* \leq s^*$.
- If $Rbcd$ and $a \leq b$, then $Racd$

Note that worldly situations play no role in this model theory. We can add them. In fact, we can replace the class of logical situations in the specification of a frame with the class of worldly situations and specify that for each worldly situation s , if Cst , then $t \leq s$.

B.2 MODELS

A model for R is a pair $\langle \mathfrak{F}, v \rangle$ where $\mathfrak{F}$ is an R-frame and v is a value assignment from propositional variables into Sit^2 such that it satisfies *hereditariness*, that is, if $s \in v(p)$ and $s \leq t$, then $t \in v(p)$. Each value assignment v determines an interpretation relation $\models_v$ such that:

- $s \models_v p$ iff $s \in v(p)$;
- $s \models_v A \wedge B$ iff $s \models_v A$ and $s \models_v B$;

- $s \models_v A \vee B$ iff $s \models_v A$ or $s \models_v B$;
- $s \models_v A \wedge B$ iff $\forall x(Csx \supset x \models_v A)$;
- $s \models_v A \rightarrow B$ iff $\forall x \forall y((Rxy \ \& \ x \models_v A) \supset y \models_v B)$.

B.3 ADDING MODALITY

B.3.1 THE BASE LOGIC

First, we must add the unary propositional operator $\Box$ to our language. The formation rules are extended in the usual way. Our base system – the logic RC – is the following system:

B.3.2 DEFINED OPERATOR

$$\Diamond A =_{df} \sim \Box \sim A$$

B.3.3 AXIOMS

R All substitution instances of theorems of R;

C $(\Box A \wedge \Box B) \rightarrow \Box(A \wedge B)$.

DV $\Box(A \vee B) \rightarrow (\Box A \vee \Diamond B)$.

B.3.4 RULES

The rules are MP, $\wedge I$ and

$$\frac{\vdash A \rightarrow B}{\therefore \vdash \Box A \rightarrow \Box B} \text{ RM}$$

B.4 SEMANTICS FOR RC AND ITS EXTENSIONS

In this section I rely heavily on (Fuhrmann 1990) and (Mares 1994).

B.4.1 FRAME THEORY

An RC-frame is a quintuple $\langle Sit, Logical, R, M, * \rangle$ such that $\langle Sit, Logical, R, * \rangle$ is an R-frame and

- if Mst , $M s^* t^*$;
- if $s \sqsubseteq s'$ and $M s' t'$, then there is a $t \sqsubseteq t'$ such that Mst .

B.4.2 THEORY OF TRUTH

An RC-model is a sextuple $\langle Sit, Logical, R, M, *, v \rangle$ such that $\langle Sit, Logical, R, M, * \rangle$ is an RC-frame, v satisfies hereditariness, and v determines a satisfaction relation $\models$ such that $\models$ obeys all the truth clauses for R-models

and

- $s \models \Box A$ if and only if $\forall x (Mst \supset x \models_v A)$.

B.5 CORRELATIONS

(N) $\frac{\vdash A}{\therefore \vdash \Box A}$ is made valid by the postulate:

$$\forall x \forall y ((x \in \text{Worldly} \ \& \ Mxy) \supset y \in \text{Worldly})$$

Name	Axiom scheme	Semantic postulate
K $\rightarrow$	$\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$	hume
D	$\Box A \rightarrow \Diamond A$	$\exists x Msx$
4	$\Box A \rightarrow \Box \Box A$	$\exists x (Msx \ \& \ Mxt) \supset Mst$
T	$\Box A \rightarrow A$	Mss
B	$A \rightarrow \Box \Diamond A$	$Mst \supset Mts$

B.6 QUANTIFICATION

We change our language to include predicates (of any arity), quantifiers ($\forall$ and $\exists$), individual variables, and individual constants.

The axiom schemes for the logic RQ are the following:

1. $\forall x A \rightarrow A[c/x]$
2. $\forall x (A \rightarrow B) \rightarrow (\forall x A \rightarrow \forall x B)$
3. $\forall x (A \vee B) \rightarrow (A \vee \forall x B)$, where x is not free in A .

The additional rule is the rule of universal generalisation, viz.,

$$\frac{\vdash A}{\vdash \forall x A}.$$

The semantics for quantified relevant logic were developed by Kit Fine ((Fine 1988) and (Anderson *et al.* 1992), §53). A quantificational R-model is a structure $\langle \mathcal{F}, \mathcal{D}, D, Q, \rightarrow, v \rangle$, where $\mathcal{F}$ is a set of R-frames, $\mathcal{D}$ is a set of sets of individuals (domains), D is a function from situations to domains, Q is a binary relation on situations, $\rightarrow$ is a partial function from situations and pairs of individuals to situations ($\overset{i,j}{\rightarrow} s$ is defined if and only if i and j are in $D(s)$), and v is a value assignment, which satisfy the postulates given below. Following Fine, I use ' s^+ ' to designate a situation such that Qss^+ and ' $\overset{i,j}{\rightarrow} s$ ' to designate ' $\overset{i,j}{\rightarrow} s$ ' where the i and j are arbitrary.

The following are Fine's postulates on quantificational frames, modified slightly to fit my notation:

1. if s and t are in the same frame, then $D(s) = D(t)$;
2. if Qst , then $D(s) \subseteq D(t)$;
3. if $\alpha \subseteq D(t)$, then there is an s such that $D(s) = \alpha$ and Qst ;

4. if Qst and $Qs't$, where $D(s) = D(s')$, then $s = s'$ (where Qst , we call s , ' $t \downarrow$ ');
5. if Qst and $D(s) = D(t)$, then $s = t$;
6. if Qst and Qsu and $D(t) \cap D(u) = D(s)$, then there are situations x and y such that $D(x) = D(y) = D(t) \cup D(u)$, Qtx , Quy , and $x \leq y$;
7. Q is transitive and reflexive;
8. if $s \leq t \downarrow$, then there is an s^+ such that $s^+ \leq t$;
9. $s^* \downarrow = s \downarrow^*$;
10. if $Rstu$, then $Rs \downarrow t \downarrow u \downarrow$;
11. if $Rstu$, then for every s^+ there are t^+ and u^+ such that $Rs^+t^+u^+$;
12. if $Rstu$, then for every u^+ there are s^+ and t^+ such that $Rs^+t^+u^+$;
13. if s is logical, then for each α such that $D(s) \subseteq \alpha$, there is a logical s^+ in α ;
14. $s \leq \vec{s}$;
15. if $s \leq t$, then $\vec{s} \leq \vec{t}$;
16. $\vec{\vec{s}} = \vec{s}$;
17. $\vec{\vec{s}}^* = \vec{s}^*$;
18. if $s \downarrow = s \downarrow$, then $s = \overset{i,j \rightarrow}{s}$;
19. if $Rstu$, then $R\vec{s} \vec{t} \vec{u}$;
20. if $Rstu$ and $s = \overset{i,j \rightarrow}{s}$, then there is some situation $u' \leq u$ such that $u' = \overset{i,j \rightarrow}{u'}$ and $Rstu'$;
21. if $D(s) \subset \alpha$, $i \in D(s)$ and $j \in \alpha \setminus D(s)$, then there is some situation t in β such that Qst and $t = \overset{i,j \rightarrow}{t}$.

The following are constraints on value assignments:

1. if $\langle v(a_1), \dots, v(a_n) \rangle \in v(P, s)$ and Qst , then $\langle v(a_1), \dots, v(a_n) \rangle \in v(P, t)$;
2. if $v(a_1), \dots, v(a_n) \in D(s)$, Qst , and $\langle v(a_1), \dots, v(a_n) \rangle \in v(P, t)$, then $\langle v(a_1), \dots, v(a_n) \rangle \in v(P, s)$;
3. if $\langle i_1, \dots, i_n \rangle \in v(P, s)$, then $\langle i'_1, \dots, i'_n \rangle \in v(P, \overset{j,k \rightarrow}{s})$, where $i'_p = i_p$ if $i_p \neq j$ and $i_p \neq k$ and if $i_p = j$ or $i_p = k$, then i'_p is either j or k , for $1 \leq p \leq n$.

The truth condition for the universal quantifier is

$$s \models_v \forall x A \text{ iff } \exists \alpha (D(s) \subset \alpha \ \& \ \exists \mu (\mu(x) \in (\alpha \setminus D(s)) \ \& \ \mu \text{ is an } x\text{-variant of } v \ \& \ \forall t (Qst \supset t \models_\mu A))).$$

We can then derive Fine's so-called 'orthodox clause', viz.,

$$s \models_v \forall x A \text{ iff } \forall t (Qst \ \& \ \mu(x) \in D(t)) \supset t \models_\mu A,$$

for every μ , an x -variant of v . The orthodox clause is the one discussed in chapter 6 above.

B.7 IDENTITY

An R-identity model is a structure $\langle \mathcal{F}, \mathcal{D}, D, Q, \rightarrow, \approx, v \rangle$, where $\langle \mathcal{F}, \mathcal{D}, D, Q, \rightarrow, v \rangle$ is a quantificational R-model, $\approx$ is a relation between situations and ordered pairs of individuals such that if $(i, j) \in \approx_s$, then i and j are in $D(s)$ and which is such that

1. if $(i, j) \in \approx_s$, then $(j, i) \in \approx_s$;
2. if $(i, j) \in \approx_s$ and $(j, k) \in \approx_s$, then $(i, k) \in \approx_s$;
3. if $s \trianglelefteq t$, then $\approx_s \subseteq \approx_t$;
4. if $(i, j) \in \approx_s$, then, for all t such that Qst , $(i, j) \in \approx_t$;
5. if $i, j \in D(s \downarrow)$ and $(i, j) \in \approx_s$, then $(i, j) \in \approx_{s \downarrow}$;
6. if $(j, k) \in \approx_{i, j \rightarrow}$, then $(i, k) \in \approx_{i, j \rightarrow}$;
7. if $s \in \text{Logical}$ and $i \in D(s)$, then $(i, i) \in \approx_s$;
8. if $(i, j) \in \approx_s$ and $(j, k) \in \approx_{s^*}$, then $(i, k) \in \approx_{s^*}$;
9. if $(i, j) \in \approx_s$, then s^* is weakly symmetrical in i, j ;
10. if $\langle i_1, \dots, i_n \rangle \in v(P, s)$ and $(i, j) \in \approx_s$, then where $i'_p = i_p$ if $i_p \neq j$ and $i'_p \neq k$ if $i_p = j$ or $i_p = k$, then i'_p is either j or k , for $1 \leq p \leq n$.

The truth condition for identity is the obvious one, i.e., $s \models_v a = b$ if and only if $(v(a), v(b)) \in \approx_s$.

11. if $\langle i_1, \dots, i_n \rangle \in v(P, s)$ and $(i, j) \in \approx_s$, then $\langle i'_1, \dots, i'_n \rangle \in v(P, s^*)$, where $i'_p = i_p$ and if $i_p \neq j$ and $i_p \neq k$ and if $i_p = j$ or $i_p = k$, then i'_p is either j or k , for $1 \leq p \leq n$.

The truth condition for identity is the obvious one, i.e., $s \models_v a = b$ if and only if $(v(a), v(b)) \in \approx_s$.

B.8 CONDITIONALS

A conditional frame is a structure $\langle \text{Sit}, \text{Logical}, \text{Poss}, R, C, \text{Prop} \rangle$ where $\langle \text{Sit}, \text{Logical}, \text{Poss}, R \rangle$ is an R-frame, Prop is a non-empty set of subsets of Sit (propositions), and C is a relation between ordered pairs of Sit and ordered triples of situations, satisfying the conditions below:

- If X and Y are in Prop , then $X \cap Y$, $X \cup Y$, $\{s : \forall x \forall y ((Rsxy \ \& \ x \in X) \supset y \in Y)\}$, and $\{s : \forall x \forall y ((C_{XY}sxy \ \& \ x \in X) \supset y \in Y)\}$ are also in Prop .
- If X is in Prop , then $\{s : \forall x (Csx \supset x \notin X)\}$ is in Prop .
- If $C_{XY}stu$, then $Rstu$.
- If $C_{X(Y \cap Z)}stu$, then $C_{XY}stu$.
- If $C_{XY}sss$.
- If $C_{XY}ss^*s^*$.

We add the connective, $\Rightarrow$ (the indicative conditional), to the language. The truth condition for the conditional is the following:

$$s \models_v A \Rightarrow B \text{ iff } \forall x \forall y ((C_{|A||B|}sxy \ \& \ x \models_v A) \supset y \models_v B).$$

B.9 COUNTERFACTUALS

A counterfactual frame is a structure $\langle Sit, Logical, Poss, R, C, B, Prop \rangle$ where $\langle Sit, Logical, Poss, R, C, Prop \rangle$ is a conditional frame, and B is a function from pairs of propositions and situations to sets of situations (base situations), satisfying the conditions below:

- If $s \in X$, then $s \in B(X, Y, s)$.
- If $s^* \notin Y$, then $s \in B(X, Y, s)$.

In addition, we modify $Prop$ so that it satisfies the conditions for a conditional frame but also satisfies the condition that, if X and Y are in $Prop$, $\{s : \forall x(B(X, Y, s) \supset \forall y\forall z((C_{XY}xyz \ \& \ y \in X) \supset z \in Y))\}$ is also in $Prop$.

We now add the connective, $\Box \Rightarrow$ (counterfactual conditional), to the language. The truth condition for the counterfactual is

$$s \models_v A \Box \Rightarrow B \text{ iff } \forall x(B(|A|, |B|, s) \supset \forall y\forall z((C_{|A||B|}xyz \ \& \ y \models_v A) \supset z \models_v B)).$$

Glossary

conditional A statement of the form ‘If ... then’. Natural language conditionals are usually divided into indicative conditionals and counterfactual conditionals. A *counterfactual* conditional is, roughly, a conditional that is stated even though its antecedent is believed to be false.

conjunction The logical connective that is expressed in English by the word ‘and’. In this book, we look at two sorts of conjunction: extensional conjunction ($\wedge$) discussed in chapter 2 and intensional conjunction, or *fusion* ($\circ$), discussed in chapter 9.

disjunction The logical connective that is expressed in English by the word ‘or’. In this book, we examine two sorts of disjunction: extensional disjunction ($\vee$) discussed in chapter 2 and intensional disjunction, or *fission* ($\oplus$), discussed in chapter 9.

E A relevant logic that was supposed to capture the notion of entailment. This logic is discussed in chapter 6.

entailment In this book, ‘ A entails B ’ is taken to mean ‘ A necessarily implies B ’. This is explained in chapter 6. Formally, ‘ A entails B ’ is written ‘ $A \multimap B$ ’. Entailment is discussed in chapter 6.

falsum A propositional constant, f , that is false in all possible worlds. In this book, the falsum is interpreted as meaning ‘Something impossible occurs’. See chapter 5.

Hilbert-style axiom system A way of defining a logical system. In a Hilbert-style axiom system, a logical system is identified with a set of statements (called ‘axioms’) and a set of rules. One derives theorems in this sort of system by constructing proofs that begin with one or more axioms and applying rules to them and then continuing by using the axioms again, or theorems already proved, together with the rules to derive more theorems.

implication In this book, ‘ A implies B ’ is taken to mean ‘ B contingently follows from A ’. This is explained in chapter 3. Formally, ‘ A implies B ’ is written ‘ $A \rightarrow B$ ’.

intersection An operation in set theory. The intersection of two sets X and Y (the intersection is written ‘ $X \cap Y$ ’) is the set of things that are in both X and Y . Thus, for example, the intersection of the sets $\{1, 2, 3\}$ and $\{2, 3, 4\}$ is the set $\{2, 3\}$.

modal logic Logic that includes symbols that mean ‘necessarily’ and ‘possibly’ or closely related notions. Modal logic is discussed in chapters 2 and 6.

multiset A multiset is a collection $[A_1, \dots, A_n]$ that is like a set in that the members of a multiset have no order but it is unlike a set in that a single thing can occur in a single multiset more than once.

natural deduction A form of proof theory that allows the use of hypotheses. This is one of the standard forms of proof theory that is taught to students in an introductory course in symbolic logic. Natural deduction is discussed in chapters 1 through 5.

negation The logical connective often expressed in English by 'not'.

possible world A possible world is a universe (like our own) that meets some conditions of coherence and completeness. For example, no possible world contains inconsistent information. In addition all possible worlds are complete, that is, for each proposition, p a world w either makes p or its negation true. Possible worlds are discussed in chapters 2 and 3.

proof theory A method for deriving theorems of a logic. The sorts of proof theory discussed in this book are natural deduction systems, sequent calculi, and Hilbert-style axiom systems.

quantifiers Logical particles are expressed in English by phrases such as 'for all', 'for some', 'there are', 'most' and 'many'. Quantifiers are discussed in chapter 6.

R The logic of relevant implication. This is the main logic discussed in this book.

sequent calculus A form of proof theory that allows the derivation of *sequents*. A sequent is a structure $\Gamma \vdash \Delta$, where Γ and Δ are *multisets* of formulae. A sequent represents a derivation. So, for example, the sequent $A \rightarrow B, A \vdash B$ represents the derivation of B From $A \rightarrow B$ and A .

set A collection of things that obeys the principles of a theory called 'set theory'. There are various different set theories. The set theory used in this book is non-well-founded set theory and is explained in chapter 3.

situation A situation is like a possible world in that it makes statements true or false, but unlike a possible world, a situation may not decide whether a given statement is true or false and some situations make some contradictions true.

theorem A statement that can be proven. Some theorems are theorems of a mathematical system, for example, $2 + 2 = 4$ is a theorem of arithmetic.

union An operation in set theory. The union of two sets X and Y (the union is written ' $X \cup Y$ ') is the set of things that are in either X or in Y . Thus, for example, the intersection of the sets $\{1, 3, 5\}$ and $\{1, 2, 4, 6\}$ is the set $\{1, 2, 3, 4, 5, 6\}$.

References

- (Ackermann 1956) Wilhelm Ackermann. Begründung einer strengen Implikation. *The Journal of Symbolic Logic*, 21: 113–28, 1956.
- (Adams 1975) Ernest W. Adams. *The Logic of Conditionals*. Reidel, Dordrecht, 1975.
- (Anderson and Belnap 1975) Alan R. Anderson and Nuel D. Belnap. *Entailment. Logic of Relevance and Necessity*, vol. 1. Princeton University Press, Princeton, NJ, 1975.
- (Anderson *et al.* 1992) Alan R. Anderson, Nuel D. Belnap and J. Michael Dunn. *Entailment. Logic of Relevance and Necessity*, vol. 2. Princeton University Press, Princeton, NJ, 1992.
- (Armstrong 1983) David M. Armstrong. *What is a Law of Nature?* Cambridge University Press, Cambridge, 1983.
- (Armstrong 1997) David M. Armstrong. *A World of States of Affairs*. Cambridge University Press, Cambridge, 1997.
- (Barwise and Etchemendy 1987) Jon Barwise and John Etchemendy. *The Liar. An Essay on Truth and Circularity*. Oxford University Press, Oxford, 1987.
- (Barwise and Moss 1996) Jon Barwise and Lawrence Moss. *Vicious Circles*. CSLI, Stanford, 1996.
- (Barwise and Perry 1983) Jon Barwise and John Perry. *Situations and Attitudes*. MIT Press, Cambridge, MA, 1983.
- (Barwise and Seligman 1997) Jon Barwise and Jerry Seligman. *Information Flow. The Logic of Distributed Systems*. Cambridge University Press, Cambridge, 1997.
- (Barwise 1989) Jon Barwise. *The Situation in Logic*. CSLI, Stanford, 1989.
- (Barwise 1993) Jon Barwise. Constraints, channels, and the flow of information. In P. Aczel, Y. Katagiri and S. Peters, eds., *Situation Theory and its Applications*, vol. 3. CSLI, Stanford, 1993.
- (Belnap 1977a) Nuel D. Belnap. How a computer should think. In *Contemporary Aspects of Philosophy*, pp. 30–55. Oriel Press, Stocksfield, 1977.
- (Belnap 1977b) Nuel D. Belnap. A useful four-valued logic. In *Modern Uses of Multiple-Valued Logic*, pp. 8–37. D. Reidel, Dordrecht, 1977.
- (Bennett 1984) Jonathan Bennett. Counterfactuals and temporal direction. *Philosophical Review*, 93: 57–91, 1984.
- (Boolos 1993) George Boolos. *The Logic of Provability*. Cambridge University Press, Cambridge, 1993.

- (Brady 2003) Ross Brady. *Relevant Logics and their Rivals, Volume II*. Ashgate, Aldershot, 2003. With contributions by Martin Bunder, André Fuhrmann, Andréa Loparić, Edwin D. Mares, Robert K. Meyer and Alasdair Urquhart.
- (Brady forthcoming) Ross Brady. *Universal Logic*. CSLI, Stanford, forthcoming.
- (Brandom 1994) Robert B. Brandom. *Making it Explicit*. Harvard University Press, Cambridge, MA, 1994.
- (Brandom 2000) Robert B. Brandom. *Articulating Reasons*. Harvard University Press, Cambridge, MA, 2000.
- (Burgess 1981) John P. Burgess. Relevance: A fallacy? *Notre Dame Journal of Formal Logic*, 22: 97–104, 1981.
- (Carroll 1994) John W. Carroll. *Laws of Nature*. Cambridge University Press, Cambridge, 1994.
- (Chellas 1974) Brian F. Chellas. Conditional obligation. In *Logical Theory and Semantic Analysis: Essays Dedicated to Stig Kanger on his Fiftieth Birthday*. Reidel, Dordrecht, 1974.
- (Chellas 1980) Brian F. Chellas. *Modal Logic*. Cambridge University Press, Cambridge, 1980.
- (Cooper 1993) Robin Cooper. Generalized quantifiers and resource situations. In *Situation Theory and its Applications*, vol. 3, pp. 191–211. CSLI, Stanford, 1993.
- (Copeland 1979) B. J. Copeland. On when a semantics is not a semantics: Some reasons for disliking the Routley–Meyer semantics for relevance logic. *Journal of Philosophical Logic*, 8: 399–413, 1979.
- (Copeland 1983) B. J. Copeland. Pure semantics and applied semantics: a rejoinder to Routley, Routley, Meyer and Martin: ‘On the philosophical basis of relevant logic semantics’. *Topoi*, 2: 192–204, 1983.
- (Cresswell 1967) M. J. Cresswell. The interpretation of some Lewis systems of modal logic. *Australasian Journal of Philosophy*, 45: 198–206, 1967.
- (Cresswell 1990) M. J. Cresswell. *Entities and Indices*. Kluwer, Dordrecht, 1990.
- (Devlin 1991) Keith Devlin. *Logic and Information*. Cambridge University Press, Cambridge, 1991.
- (Dowty et al. 1985) David Dowty, Robert E. Wall and Stanley Peters. *Introduction to Montague Semantics*. Reidel, Dordrecht, 1985.
- (Dretske 1977) Fred Dretske. Laws of nature. *Philosophy of Science*, 44, 1977.
- (Dunn 1976) J. Michael Dunn. Intuitive semantics for first-degree entailments and ‘coupled trees’. *Philosophical Studies*, 29: 149–68, 1976.
- (Dunn 1986) J. Michael Dunn. Relevance logic and entailment. In F. Guentner and D. Gabbay, ed., *Handbook of Philosophical Logic*, vol. 3, pp. 117–224. Reidel, Dordrecht, 1986.
- (Dunn 1987) J. Michael Dunn. Relevant predication I: The formal theory. *Journal of Philosophical Logic*, 16: 347–81, 1987.
- (Dunn 1990) J. Michael Dunn. Relevant predication 3: Essential properties. In J. M. Dunn and A. Gupta, eds., *Truth or Consequences: Essays in Honour of Nuel Belnap*, pp. 77–95. Kluwer, Dordrecht, 1990.
- (Dunn 1993) J. Michael Dunn. Star and perp. *Philosophical Perspectives*, 7: 331–57, 1993.

- (Field 1989) Hartry Field. *Realism, Mathematics and Modality*. Blackwell, Oxford, 1989.
- (Fine 1974) Kit Fine. Models for entailment. *Journal of Philosophical Logic*, 3: 347–72, 1974.
- (Fine 1988) Kit Fine. Semantics for quantified relevance logic. *Journal of Philosophical Logic*, 17: 27–59, 1988.
- (Fitch 1952) Frederic Fitch. *Symbolic Logic*. Ronald Press, New York, 1952.
- (Forbes 1985) Graeme Forbes. *The Metaphysics of Modality*. Oxford University Press, Oxford, 1985.
- (Fraenkel *et al.* 1973) A. A. Fraenkel, Y. Bar-Hillel and A. Levy. *Foundations of Set Theory*. North Holland, 1973.
- (Friedman and Meyer 1992) Harvey Friedman and Robert K. Meyer. Whither relevant arithmetic. *The Journal of Symbolic Logic*, 57: 824–31, 1992.
- (Fuhrmann 1990) André Fuhrmann. Models for relevant modal logics. *Studia Logica*, 49, 1990.
- (Gabbay 1972) Dov M. Gabbay. A general theory of the conditional in terms of a ternary operator. *Theoria*, 38: 97–104, 1972.
- (Gabbay 1976) Dov M. Gabbay. *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*. Reidel, Dordrecht, 1976.
- (Garfield 1990) Jay Garfield. The dog: relevance and rationality. In J. M. Dunn and Anil Gupta, eds., *Truth or Consequences*, pp. 97–109. Kluwer, Dordrecht, 1990.
- (Gentzen 1969) Gerhard Gentzen. Investigations into logical deduction. In M. E. Szabo, ed., *The Collected Papers of Gerhard Gentzen*, pp. 70–131. North Holland, Amsterdam, 1969.
- (Goble 1999) Lou Goble. Deontic logic with relevance. In P. McNamara and H. Prakken, eds., *Norms, Logics, and Information Systems*, pp. 331–45. IOS Press, Amsterdam, 1999.
- (Goldblatt 1974) Robert Goldblatt. Semantic analysis of orthologic. *Journal of Philosophical Logic*, 3: 19–35, 1974. Reprinted in Goldblatt, *Mathematics of Modality*, CSLI, Stanford, 1993, ch. 2.
- (Grafton 1989) Sue Grafton. *F is for Fugitive*. Pan, London, 1989.
- (Grice 1989) H. Paul Grice. *Logic and Conversation*, pp. 1–137. Harvard University Press, Cambridge, MA, 1989.
- (Halpin 1991) John Halpin. The miraculous conception of counterfactuals. *Philosophical Studies*, 63: 271–90, 1991.
- (Hansson 1969) Bengt Hansson. An analysis of some deontic logics. *Noûs*, 3: 373–89, 1969.
- (Hare 1970) R. M. Hare. Meaning and speech acts. *Philosophical Review*, 79: 3–24, 1970.
- (Hazen 1979) Allen Hazen. Counterpart theoretic semantics for modal logic. *Journal of Philosophy*, 76: 319–38, 1979.
- (Horn 1989) Laurence Horn. *A Natural History of Negation*. Chicago University Press, Chicago, 1989.

- (Hughes and Cresswell 1968) G. E. Hughes and M. J. Cresswell. *An Introduction to Modal Logic*. Methuen, 1968.
- (Hughes and Cresswell 1996) G. E. Hughes and M. J. Cresswell. *A New Introduction to Modal Logic*. Routledge, London, 1996.
- (Israel and Perry 1990) David Israel and John Perry. What is information? In P. P. Hanson, ed., *Information, Language, and Cognition*, pp. 1–19. University of British Columbia Press, Vancouver, 1990.
- (Jackson 1987) Frank Jackson. *Conditionals*. Blackwell: Oxford, 1987.
- (Jeffreys 1973) Harold Jeffreys. *Scientific Inference*. Cambridge University Press, Cambridge, 1973.
- (Joyce 2001) Richard Joyce. *The Myth of Morality*. Cambridge University Press, Cambridge, 2001.
- (Kratzer 1977) Agelika Kratzer. What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy*, 1: 337–55, 1977.
- (Kreisel 1967) Georg Kreisel. Mathematical logic: what has it done for the philosophy of mathematics? In Ralph Schoenman, ed., *Bertrand Russell: Philosopher of the Century*, pp. 201–72. Little, Brown and Company, Boston, 1967.
- (Kremer 1999) Philip Kremer. Relevant identity. *Journal of Philosophical Logic*, 28: 199–22, 1999.
- (Kripke 1965a) Saul A. Kripke. Semantical analysis of intuitionist logic – I. In J. N. Crossley and Michael Dummett, eds., *Formal Systems and Recursive Functions*, pp. 92–129. North Holland, Amsterdam, 1965.
- (Kripke 1965b) Saul A. Kripke. Semantical analysis of modal logic – II. In J. Addison, L. Henkin and A. Tarski, eds., *The Theory of Models*, pp. 206–20. North Holland, Amsterdam, 1965.
- (Kripke 1972) Saul A. Kripke. *Naming and Necessity*. Harvard University Press, Cambridge, MA, 1972.
- (Lance and Kremer 1996) Mark Lance and Philip Kremer. The logical structure of linguistic commitment II: relevant commitment entailment. *Journal of Philosophical Logic*, 25: 425–49, 1996.
- (Lance 1988) Mark Lance. The logic of contingent relevant implication: a conceptual incoherence in the intuitive foundation of R. *Notre Dame Journal of Formal Logic*, 29: 520–9, 1988.
- (Lange 1993) Marc Lange. When would natural laws have been broken? *Analysis*, 53: 262–9, 1993.
- (Lemmon 1965) E. J. Lemmon. *Beginning Logic*. Nelson, London, 1965.
- (Levi 1996) Isaac Levi. *For the Sake of Argument*. Cambridge University Press, Cambridge, 1996.
- (Lewis and Langford 1959) C. I. Lewis and C. H. Langford. *Symbolic Logic*. Dover, New York, 1959.
- (Lewis 1917) C. I. Lewis. The issues concerning material implication. *Journal of Philosophy, Psychology, and Scientific Methods*, 14: 350–6, 1917.
- (Lewis 1968) David K. Lewis. Counterpart theory and quantified modal logic. *Journal of Philosophy*, 65: 113–26, 1968.

- (Lewis 1973) David K. Lewis. *Counterfactuals*. Harvard University Press, Cambridge, MA, 1973.
- (Lewis 1979) David K. Lewis. Counterfactual dependence and time's arrow. *Noûs*, 13: 32, 1979.
- (Lewis 1984) David K. Lewis. Postscripts to 'truth in fiction'. In *Philosophical Papers*, vol. 1, pp. 276–80, 1984.
- (Lewis 1986) David K. Lewis. *On the Plurality of Worlds*. Oxford University Press, Oxford, 1986.
- (Lycan and Shapiro 1986) William G. Lycan and Stuart Shapiro. Actuality and essence. In *Midwest Studies in Philosophy XI: Studies in Essentialism*. University of Minnesota Press, Minneapolis.
- (Lycan 1993) William G. Lycan. MMP, RIP. In *Philosophical Perspectives*, vol. 7, pp. 411–27. Ridgeview, Atascadero, 1993.
- (Maksimova 1971) Larisa L. Maksimova. An interpretation and separation theorem for the logical systems E and R. *Algebra and Logic*, 7: 102–16, 1971.
- (Mares and Fuhrmann 1995) Edwin D. Mares and André Fuhrmann. A relevant theory of conditionals. *Journal of Philosophical Logic*, 24: 645–65, 1995.
- (Mares and Meyer 2001) Edwin D. Mares and Robert K. Meyer. Relevant logics. In Louis Goble, ed., *Blackwell's Guide to Philosophical Logic*. Blackwell, Oxford, 2001.
- (Mares 1992) Edwin D. Mares. Semantics for relevance logic with identity. *Studia Logica*, 51: 1–20, 1992.
- (Mares 1994) Edwin D. Mares. Mostly Meyer model models. *Logique et analyse*, 146: 119–28, 1994.
- (Mares 1995) Edwin D. Mares. A star-free semantics for R. *The Journal of Symbolic Logic*, 60: 579–90, 1995.
- (Mares 1996) Edwin D. Mares. Relevant logic and the theory of information. *Synthese*, 109: 345–60, 1996.
- (Mares 2000) Edwin D. Mares. Even dialetheists should hate contradictions. *Australasian Journal of Philosophy*, 78: 503–16, 2000.
- (Mares 2002a) Edwin D. Mares. A paraconsistent theory of belief revision. *Erkenntnis*, 56: 229–46, 2002.
- (Mares 2002b) Edwin D. Mares. Relevance logic. In Dale Jacquette, ed., *Companion to Philosophical Logic*, pp. 609–27. Blackwell, Oxford, 2002.
- (Mares forthcoming) Edwin D. Mares. Four-Valued Semantics for the Relevant Logic R. *Journal of Philosophical Logic* (forthcoming).
- (Mares forthcoming) Edwin D. Mares. Relevant implication and the indicative conditional. *Synthese* (forthcoming).
- (Martin and Meyer 1982) Errol P. Martin and Robert K. Meyer. Solution to the P–W problem. *The Journal of Symbolic Logic*, 47: 869–86, 1982.
- (McGee 1985) Van McGee. A counterexample to modus ponens. *The Journal of Philosophy*, 82: 462–71, 1985.
- (Meyer and Dunn 1969) Robert K. Meyer and J. M. Dunn. E, R, and γ . *The Journal of Symbolic Logic*, 34: 460–74, 1969.

- (Meyer and Mortensen 1984) Robert K. Meyer and Chris Mortensen. Models for relevant arithmetics. *The Journal of Symbolic Logic*, 49: 917–29, 1984.
- (Meyer and Routley 1973) Robert K. Meyer and Richard Routley. Classical relevant logic I. *Studia Logica*, 32: 51–68, 1973.
- (Meyer 1971) Robert K. Meyer. Entailment. *Journal of Philosophy*, 68, 1971.
- (Meyer 1976) Robert K. Meyer. Relevant arithmetic. *Bulletin of the Section of Logic*, 5: 133–7, 1976.
- (Moore 1922) G. E. Moore. External and internal relations. In *Philosophical Studies*. Routledge and Kegan Paul, London, 1922.
- (Mortensen 1983) Chris Mortensen. The validity of disjunctive syllogism is not so easily proved. *Notre Dame Journal of Formal Logic*, 24: 35–40, 1983.
- (Mortensen 1986) Chris Mortensen. Reply to Burgess and Read. *Notre Dame Journal of Formal Logic*, 27: 195–200, 1986.
- (Mortensen 1995) Chris Mortensen. *Inconsistent Mathematics*. Kluwer, Dordrecht, 1995.
- (Paoli 2002) Francesco Paoli. *Substructural Logics: A Primer*. Kluwer, Dordrecht, 2002.
- (Parsons 1983) Charles Parsons. What is the iterative conception of set? In *Mathematics in philosophy: selected essays*, pages 268–97. Cornell University Press, Ithaca, NY, 1983.
- (Parsons 1984) Terence Parsons. Assertion, denial, and the liar paradox. *Journal of Philosophical Logic*, 13: 137–52, 1984.
- (Pollock 1984) John L. Pollock. *Foundations of Philosophical Semantics*. Princeton University Press, Princeton, 1984.
- (Price 1990) Huw Price. Why ‘not’? *Mind*, 99: 221–38, 1990.
- (Priest 1986) Graham Priest. Contradiction, belief, and rationality. *Proceedings of the Aristotelian Society*, 86: 99–116, 1986.
- (Priest 1987) Graham Priest. *In Contradiction. An Essay on the Transconsistent*. Nijhoff, 1987.
- (Priest 1990) Graham Priest. Boolean negation and all that. *Journal of Philosophical Logic*, 19: 201–15, 1990.
- (Priest 1992) Graham Priest. What is a non-normal world? *Logique et Analyse* 139–40: 291–302.
- (Priest 2000) Graham Priest. Motivations for paraconsistency: The slippery slope from classical logic to dialetheism. In Diderik Batens, Chris Mortensen, Graham Priest and Jean-Paul Van Bendegem, eds., *Frontiers of Paraconsistent Logic*. Research Studies Press, Baldock, 2000.
- (Priest 2001) Graham Priest. *An Introduction to Non-Classical Logic*. Cambridge University Press, Cambridge, 2001.
- (Prior 1976) Arthur N. Prior. The autonomy of ethics. In *Papers in Logic and Ethics*, pp. 88–96. Duckworth, London, 1976.
- (Quine 1950) W. V. O. Quine. *Methods of Logic*. Harvard University Press, Cambridge, MA, 1950.
- (Read 1983) Stephen Read. Burgess on relevance: A fallacy indeed. *Notre Dame Journal of Formal Logic*, 24: 473–81, 1983.

- (Read 1988) Stephen Read. *Relevant Logic*. Blackwell, Oxford, 1988.
- (Restall 1995) Greg Restall. Four-valued semantics for relevant logics (and some of their rivals) *Journal of Philosophical Logic* 24 (1995): 139–160.
- (Restall 1996) Greg Restall. Information flow and relevant logics. In Jerry Seligman and Dag Westerståhl, eds., *Logic, Language, and Computation, Volume I*, pp. 463–77. CSLI, Stanford, 1996.
- (Restall 2000) Greg Restall. *An Introduction to Substructural Logics*. Routledge, London, 2000.
- (Rosen 1990) Gideon Rosen. Modal fictionalism. *Mind*, 99: 327–54, 1990.
- (Routley and Meyer 1972) Richard Routley and Robert K. Meyer. The semantics of entailment III. *Journal of Philosophical Logic*, 1: 192–208, 1972.
- (Routley and Meyer 1973) Richard Routley and Robert K. Meyer. The semantics of entailment. In Hughes Leblanc, ed., *Truth, Syntax, and Modality*, pp. 199–243. North Holland, 1973.
- (Routley and Routley 1972) Richard Routley and Val Routley. Semantics for first degree entailment. *Noûs*, 6: 335–59, 1972.
- (Routley *et al.* 1982) Richard Routley, Robert K. Meyer, Val Plumwood and Ross T. Brady. *Relevant Logics and Their Rivals I*. Ridgeview, Atascadero, 1982.
- (Routley 1984) Richard Routley. The American plan completed: alternative classical-style semantics, without stars, for relevant and paraconsistent logics. *Studia Logica*, 43: 131–58, 1984.
- (Ryle 1949) Gilbert Ryle. *Concept of Mind*. Barnes and Noble, 1949.
- (Searle 1969) John Searle. *Speech Acts*. Cambridge University Press, 1969.
- (Segerberg 1971) Krister Segerberg. *An Essay in Classical Modal Logic*. University of Uppsala, Uppsala, 1971.
- (Sellars 1963) Wilfred F. Sellars. *Some Reflections on Language Games*, pp. 321–58. Routledge and Kegan Paul, London, 1963.
- (Slaney 1990) John Slaney. A general logic. *Australasian Journal of Philosophy*, 68: 74–88, 1990.
- (Stalnaker 1970) Robert Stalnaker. Probability and conditionals. *Philosophy of Science*, 37: 64–80, 1970.
- (Stalnaker 1984) Robert Stalnaker. *Inquiry*. MIT Press, Cambridge, MA, 1984.
- (Stalnaker 1986) Robert Stalnaker. Possible worlds and situations. *Journal of Philosophical Logic*, 15: 109–23, 1986.
- (Stalnaker 1999) Robert C. Stalnaker. Indicative conditionals. In *Content and Context*, pp. 63–77. Oxford University Press, Oxford, 1999. Originally published in *Philosophia* 5 (1975).
- (Tarski 1985) Alfred Tarski. The concept of truth in formalized languages. In John Corcoran, ed., *Alfred Tarski, Logic, Semantics, Metamathematics*, pp. 152–278. Hackett, Indianapolis, 1985.
- (ter Meulen 1995) Alice G. B. ter Meulen. *Representing Time in Natural Language*. MIT, Cambridge, MA, 1995.
- (Tooley 1977) Michael Tooley. The nature of laws. *Canadian Journal of Philosophy*, 7, 1977.

- (Tooley 1987) Michael Tooley. *Causation. A Realist Approach*. Oxford University Press, Oxford, 1987.
- (Urquhart 1972) Alasdair Urquhart. Semantics for relevant logics. *The Journal of Symbolic Logic*, 37: 159–69, 1972. Reprinted in Anderson, Belnap and Dunn, *Entailment*, Volume 2, 1992.
- (Urquhart 1981) Alasdair Urquhart. Intensional languages via nominalisation. *Pacific Journal of Philosophy*, 62, 1981.
- (van Bentham 1979) Johann van Bentham. What is dialectical logic? *Erkenntnis*, 14: 333–47, 1979.
- (van Fraassen 1972) Bas C. van Fraassen. The logic of conditional obligation. *Journal of Philosophical Logic*, 1: 417–38, 1972.
- (van Fraassen 1980) Bas C. van Fraassen. *The Scientific Image*. Oxford University Press, Oxford, 1980.
- (van Fraassen 1981) Bas C. van Fraassen. Essences and laws of nature. In R. Healey, ed., *Reduction, Time and Reality*. Cambridge University Press, Cambridge, 1981.
- (van Fraassen 1989) Bas C. van Fraassen. *Laws and Symmetry*. Oxford University Press, Oxford, 1989.
- (Williamson 1994) Timothy Williamson. *Vagueness*. Routledge, London, 1994.
- (Woods 1997) Michael Woods. *Conditionals*. Oxford University Press, Oxford, 1997.
- (Woollaston 1994) Lin Woollaston. Counterpart theory as a semantics for modal logic. *Logique et Analyse*, 147–8: 255–63, 1994.
- (Yagasawa 1988) Takaishi Yagasawa. Beyond possible worlds. *Philosophical Studies*, 53: 175–204, 1988.

Index

- Ackermann, Wilhelm, 96, 97, 107, 174, 175
Aczel, Peter, viii, 59, 62
Adams, Ernest, 140
 Adam's thesis, 140–2
Anderson, Alan R., 6, 11, 33–5, 37, 97, 112,
 113, 128, 171, 174, 175
anti-realism
 metaphysical, 57
 fictionalist, 57
 instrumentalist, 57
Armstrong, David, 61, 67, 68
axiom
 4, 103
 anti-foundation, 62
 of foundation, 62, 72
 T, 103
B, 201
Barwise, Jon, vii, x, 27, 39, 43, 44, 54, 59,
 62–5, 155, 156
belief revision, 192
Belnap, Nuel D. Jr., 6, 11, 33–5, 37, 87, 97,
 112, 113, 128, 163, 171, 174, 175, 177,
 200
bivalence, 27, 73, 75, 80
Brady, Ross, x, 34, 35, 198, 201–4, 206
Brandom, Robert, 20
Burgess, John, 180
channel
 information-theoretic, 54, 55
Chisholm, Roderick, 148, 149
circularity, 63, 65
Cohen, Dan, 200
compatibility, 74
compositionality, 21, 22, 24
conditional
 indicative, 144
conditionals, vii–ix, 12, 13, 15, 127,
 128
 counter-logical, 14
 counterfactual, ix, 11–14, 144–9, 152
 counterlegal, 147, 148
 counterpossible, 13, 147
 even if, 142, 143
 indicative, ix, 11–14, 125–33, 144
 material, 14
 model theory for, 137
 natural language, 11, 12
 ponable, 136
 relevant, 11, 58
 standard view of, 12–4
conjunction
 elimination, 48
 extensional, 33, 37, 49, 50, 128, 168
 intensional (see fusion), 167
 introduction, 48, 49
consistency, 183
 principle of, 80
contradiction, 82
Cooper, Robin, 40
deduction
 relevant, 35
definite description, 115
denial, 93–5, 156, 173, 184–6, 193, 195
denial operator, 94, 95
deontic logic, 108, 109, 189, 190, 207
 dyadic, 189–91
Devlin, Keith, 71
disjunction
 extensional, 33, 37, 49, 181
disjunctive syllogism, ix, 9, 81, 136, 175–80,
 185–8, 196, 206, 207
 intensional (IDS), 170, 180, 182–4
 pragmatic (PDS), 184, 185
 strong (SDS), 176, 180
domain, 21, 111, 113–15, 178, 212
 of a situation, 114–17
 of philosopher's quantification, 110
 variable, 115
Dretske, Fred, 67, 68
Duhem, Pierre, 195
Duhem–Quine problem, 195

- Dummett, Michael, 19
- Dunn, J.M., x, 28, 74, 75, 77, 79, 81, 82, 87, 89, 118, 126, 131, 141, 163, 175, 191, 192, 210
- E, 20, 97, 98, 175, 183
- entailment, 10, 11, 96–100, 102, 105
 relevant, 10–12, 98
 semantic, 30, 31, 67
- essential predication, 191
- Etchemendy, John, 62, 63
- ex falso quodlibet, 8, 9, 96, 175–7, 183, 187, 188
- fact, 62, 66
 moral, 58
- fallacy
 of modality, 106
- Fine, Kit, x, 28, 114–16, 172, 212, 213
- fission, 169, 180, 181
- Fitch, Frederick, 5, 112
- Friedman, Harvey, 196
- function, 68
- fusion, 166–71
- Gabbay, Dov, 132, 146
- Gentzen, Gerhard, 29, 34, 35, 163, 164, 169, 172
- Gibbard, Allan, 136
- GL, 100
- Goble, Lou, x, 46, 191
- Goedel, Kurt, 100
- Goldblatt, Robert, x, 74, 75
- Goodman, Nelson, 148, 149, 151, 152
- Grice, Paul, 17, 137
- Hardy, G.H., 83, 84
- Hempel, Carl, 151
- hereditariness
 is the same as persistence, 30
- Hintikka, Jaakko, 22
- identity, 83, 84, 116–18, 120, 121, 192, 214
- implication, 5–8, 10–12, 15, 26, 45, 84, 97, 99, 109, 118, 119, 127, 128, 167, 168
 and inference, 32
 and situated inference, 43–6, 50, 51
 as a relation between propositions, 45
 elimination, 5
 implication, 127
 introduction, 5, 8, 42
 intuitionist, 82
 material, 8, 36, 82, 96, 107, 126
 neighbourhood semantics for, 47
 provable, 8
 relational semantics for, 50
 relevant, 7, 10, 11, 35, 37–9, 42, 43, 45–7, 51, 53–5, 57, 58, 61, 66–8, 73, 81, 82, 84, 88, 96, 100, 104, 118, 122, 127–9
 strengthening, 97
 strict, 8, 96, 97, 99
 strict relevant, 97, 98, 100
 truth condition for, 46, 48, 51, 53, 56, 89, 105, 129
- implication relation
 on pairs of propositions and sets of situations, 46
- impossible world, 70
- inconsistency
 of situations, 75
- individual
 counterpart, 60
 trans-world, 60
- individuals, 59
- inference
 relevant, 38
 situated, 11, 41, 43–50, 55, 66, 67, 96, 98, 170, 171, 189, 207
- infin, 71
- informational link, 67, 68
- intersection
 generalised, 51
- Jackson, Frank, 141
- Joyce, Richard, 57
- Kanger, Stig, 22
- Kremer, Philip, 20, 121
- Kripke, Saul, 22, 23, 26, 27, 39, 52, 69, 74, 100, 106, 115, 120, 121
 Kripke's principle, 121
- labelled transition system, 65
- Lance, Mark, 20
- Lange, Marc, 145, 148, 150
- Langford, C.H., 9
- law of excluded middle, 80, 95, 103
- law of nature, 24, 41, 43, 44, 67, 68, 148–54
- law of non-contradiction, 80
- laws of logic, 53
- legal conservativeness, 149
- Leibniz, G.W., 23
- Levi, Isaac, 134
- Lewis, C.I., 8, 9, 22, 84, 96, 107, 176, 177
- Lewis, David, 59, 60, 68–70, 84–6, 106, 140, 152–4, 189, 190
- Locker-Lampson, Frederick, 14

- logic
 - alternative, 53
 - classical, 3, 35, 58, 74, 81, 82, 109, 165, 166, 169
 - intuitionist, 19, 35, 82, 115, 116, 165, 169
 - non-classical, vii, 74
 - relevant, vii, viii
- logic, classical, 176, 181
- Lycan, William, 105, 106, 136

- Martin, Errol, 201
- McGee, Vann, 133, 134
- McTaggart, J.M.E., 83
- metaphysics, 57
- Meyer, Robert, x, 27, 29, 32, 33, 46, 51, 52, 92, 97, 174, 175, 196–8, 201
- Meyer-Routley semantics
 - for Boolean negation, 92
- Mints, Gregor, 163
- modal logic, vii, 100, 103
 - classical, 22
 - normal, 22
 - regular, 22
- modal operator
 - truth conditions for, 100
- modality, 11, 41, 96, 101, 103
 - absolute, 69
 - alethic, 102
 - relative, 69
- model
 - intended, 67
- Montague, Richard
 - Montague grammar, vii
- Mortensen, Chris, 183, 204
- Moss, Lawrence, 59, 63–5
- multiset, 166, 167, 171, 172, 205

- natural deduction, viii, 5, 7, 8, 17
 - classical, 5, 6, 19, 34, 188
 - contraction-free, 206
 - Fitch-style, 5
 - intuitionist, 34
 - relevant, 6, 7, 11, 19, 20, 29, 30, 32–7, 42, 45, 48, 78, 83, 111, 128, 138, 164, 165, 168, 169, 171, 175, 176, 187, 199, 206, 209, 217
- necessary truth, 3, 66, 71
- necessity, 8, 10, 11, 23–5, 97, 99–102
 - absolute, 101
 - logical, 15, 99, 104
 - metaphysical, 101, 102, 105, 106, 192
 - necessity, 103
 - nomic, 24, 67, 68, 102–4
 - unrestricted, 101
- necessity, 22
- negation, 41, 61, 74–9, 81, 82, 84–95, 167–9
 - Boolean, 92, 93, 95
 - definition of in terms of impossibility
 - constant, 81
 - denial, 93
 - four-valued treatment of, 87
 - non-classical, 73
 - truth condition for, 80, 82
- neighbourhood semantics, 46, 49
- non-normal worlds, 53
- non-sequitur, 3, 4, 6, 16
- NR, 97, 98

- ontological commitment, 58, 60

- Paoli, Francesco, viii
- paradoxes
 - hypergame, 63, 64
 - of material implication, 8
 - of strict implication, 8
 - positive, 9, 10, 167
- Parwitz, Dag, 35
- Perry, John, vii, x, 27, 39, 44, 122
- persistence, 30, 31
- Plato, 76, 77
- possibility, 22, 101
 - metaphysical, 106
- possible world, vii, viii, 13, 16, 19, 22–9, 39–42, 49, 53, 57–61, 66, 69, 71, 95, 99, 101, 102, 105, 106, 114, 120, 155, 178, 189, 190, 192, 217
- Prawitz, Dag, 34
- premises, ix, 3, 4, 6, 165
 - and fusion, 169, 171, 187
 - as hypotheses in natural deduction, 5
 - bunched in multisets, 166, 167
 - bunched in sets, 165, 166
 - connection between content of and conclusion, 4, 6, 17
 - deductions with arbitrary numbers of, 31
 - discharging, 7
 - implicational, 12
 - in classical and intuitionist logic, 165
 - in sequents, 164
 - in weak relevant logic, 206
 - indexed in natural deduction proofs, 6
 - irrelevant, 6, 48
 - order of, 35
 - real use of, 6, 19, 207
- Price, Huw, 76, 77
- Priest, Graham, viii, x, 38, 53, 54, 73, 90, 91, 203, 204
 - on logic fictions, 53

- Prior, Arthur, 22
 property, 60, 61
 as function, 60
 higher-order, 60
 proposition
 implicational, 66
 Putnam, Hilary, 69, 106
- quantifier, 109, 111, 113–5, 203, 212, 217
 existential, 111–13
 generalised, 110
 natural deduction rules for, 111
 philosopher's, 110
 restricted, 110
 universal, 111, 112, 115, 116, 213
 Quine, W.V.O., 58, 100, 195
- R, 20, 35, 37–9, 54, 81, 87, 97, 98, 103, 164,
 166–8, 175, 183, 200, 202–4, 206, 210,
 211, 217
- Read, Stephen, 170, 171, 180, 181, 183, 202,
 205, 206
- realism
 ersatz, 68, 69
 ersatz (with regard to possible worlds), 59
 metaphysical, 57
 vertibrate, 68
 vertibrate (with regard to possible worlds),
 59
- reduction, 58
- reference
 direct, 69
- Reichenbach, Hans, 151
- relation, 60, 61
 accessibility, 23, 24, 26, 27, 30–3, 51, 71,
 78, 81, 89, 100–6, 115, 122, 129
 consequence, 171, 172
 on individuals as a function, 60
 part-whole (on situations), 30, 31, 52, 65
- relevant predication, 118, 191, 192
- Restall, Greg, viii, x, 54
- Routley-Meyer Semantics, ix, 27–9, 32, 36, 38,
 39, 51–4, 74, 89, 122, 188, 191, 210–12
- Russell, Bertrand, 83, 110
 Russell's paradox, 203
- S (is for “syllogism”), 200, 201
 S2, 96
 S4, 99, 103
 S5, 99
- Sellars, Wilfred, 20
- semantics, 57
- sequent, 163–7, 169
- set theory
 Bernays-von Neumann-Goedel, 62
 non-well-founded, viii, 59, 62, 63, 65, 71
 Zermelo-Fraenkel, 59, 62
- Shapiro, Stuart, 105, 106
- situation, vii, 6, 11, 16, 26–8, 30–3, 35, 36,
 39–47, 49–54, 58, 61, 65–8, 70, 71,
 73–82, 84, 86, 87, 89, 92, 94, 95, 100,
 101, 115, 144
 abstract, 41, 65, 67
 base, 145–52, 154–6, 159, 190, 191, 215
 concrete, 41
 evidential, 39
 impossible, 27, 70, 73, 75, 81, 82
 inconsistent, 84–7, 93, 95
 logical, 30, 31, 36, 37, 52, 53, 66, 79
 maximal, 78, 79
 non-normal, 52–4
 normal, 54
 partial, 40, 73, 75, 86
 possible, 66, 69, 81, 95, 155
 resource, 40
 semantics, vii, 39, 41, 61
 type, 44
 worldly, 79–81, 95
- Slaney, John, 206
- Stalnaker, Robert, 13, 15, 138
 Stalnaker's thesis, 140
- state of affairs (SOA), 61, 62, 65–8, 70, 103,
 105, 106, 120
 law-like, 67, 103
- Tarski, Alfred, 20, 21, 24, 86, 95, 171
- theory of meaning
 inferential, 20
 truth conditional, 19, 20, 22, 24
- Tooley, Michael, 67, 68
- Urelement, 62
- Urquhart, Alasdair, x, 28, 100
- validity
 classical, 3, 4
 of formulae in the Routley-Meyer
 semantics, 29, 30, 52
 relevant, 6
- van Fraassen, Bas, 151, 152
- Williamson, Timothy, 84, 85
- Yagisawa, Takashi, 54
- Zwicker, William, 63