Frank Veltman

LOGICS FOR CONDITIONALS

L O G I C S
F O R
C O N D I T I O N A L S

L O G I C S
F O R
C O N D I T I O N A L S

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van
doctor in de Wijsbegeerte
aan de Universiteit van Amsterdam,
op gezag van de Rector Magnificus
dr. D.W. Bresters,
hoogleraar in de Faculteit
der Wiskunde en Natuurwetenschappen,
in het openbaar te verdedigen
in de Aula van de Universiteit
(tijdelijk Roetersstraat 15),
op vrijdag 4 oktober 1985
te 14.30 uur

door

FRANCISCUS JACOBUS MARIA MARTINUS VELTMAN
geboren te Eindhoven

Promotores:

prof. dr. J.F.A.K. van Benthem, Rijksuniversiteit, Groningen

prof. dr. J.A.W. Kamp, University of Texas, Austin

ACKNOWLEDGEMENTS

I do not know to whom I am most indebted, to Johan van Benthem or to Hans Kamp. What I do know is that this would have been an altogether different thesis if I had not had the privilege of having *both* of them as supervisors. With his gift for anticipation Hans is the ideal person to air ones views to - even long before they deserve it. Much of what is new in the following was inspired by discussions with him. But where Hans was the one who most encouraged me to follow up my own ideas, Johan saw to it that I did not get too wild about them. He has always urged me to put things in a broader perspective and his comments and questions determined the final shape of this book.

I am grateful also to my teachers at the University of Utrecht, my former colleagues at the Erasmus University in Rotterdam, and my present colleagues at the University of Amsterdam. In this connection I want to mention in particular Jeroen van Rijen, Dick de Jongh, Roel de Vrijer, Jeroen Groenendijk, Martin Stokhof and Fred Landman: Jeroen van Rijen's ideas on the philosophy of logic greatly influenced mine; the proofs in this book have benefited much from Dick de Jongh's mathematical taste; Roel de Vrijer carefully read the manuscript and suggested several important improvements on the presentation; Jeroen Groenendijk's and Martin Stokhof's preoccupation with the role of background information in matters of semantics and pragmatics has been of great importance for the development of my own views; and Fred Landman gets the prize for asking the nastiest questions.

As will become clear from the notes many other people have by their suggestions and criticism contributed to the contents of this book. I cannot here mention each of them by name, but I am grateful to all of them.

As for editional matters, I am greatly indebted to Michael Morreau who tried his utmost to turn my double Dutch into readable English. Marjorie Pigge typed this book and I want to thank her for the great care with which she did so.

This book has been nearing completion for quite some time. Now that it is finally finished, I would like to dedicate it to all those who are relieved that it is.

Amsterdam
August, 1985

Frank Veltman

CONTENTS

PART I

THE PROBLEM OF CONDITIONALS

I.1. METHODOLOGICAL REMARKS

To those who believe that there is such a thing as the logic of conditionals this dissertation may appear to be yet another attempt to unravel its secrets - which it is not. As a logician, you can do no more than devise a logic for conditionals and try to persuade your readers to adopt it. You may succeed in doing so if you are able to demonstrate that the one you propose is a better logic for conditionals than the ones proposed so far. The phrase 'better for conditionals' should, however, not be misunderstood. In particular, it should not be interpreted as meaning 'more like the real one'. The best logic for conditionals one might propose is not that which they actually possess. It cannot be, not because this actual logic would be not good enough, but simply because there is no such thing. Whether a given logic is better than some alternative has little to do with its better fitting the facts; it is more a question of efficacy.

This is one of the theses defended in the following introductory pages. It is put forward when the question is discussed as to how one should choose between rival logical theories. That is a very natural question to ask in an introduction to the problem of conditionals. If only because so many theories have been put forward, all purporting to solve the problem, that putting yet another one on the market might seem to only add to the difficulties.

I.1.1. The case of the marbles

Here is an example which will return regularly in much of the following. There are three marbles: one red, one blue, and one yellow. They are known to be distributed among two matchboxes, called 1 and 2. The only other thing which you are told is that there is at least one marble in each box.

The various possibilities which this leaves open can be summarized as follows:

	box 1	box 2
I	blue	yellow, red
II	yellow	blue, red
III	red	blue, yellow
IV	yellow, red	blue
V	red, blue	yellow
VI	blue, yellow	red

Bearing these possibilities in mind, you will have to agree that

(1) *The yellow marble is in box 1, if both of the others are to be found in box 2*

Now suppose someone no better informed than yourself were to claim that

(2) *The yellow marble is in box 1 if the blue marble is in box 2*

You will disagree. You may even go as far as to assert

(3) *It is not so that if the blue marble is in box 2, the yellow one is in box 1*

After all, the yellow marble could just as well be in box 2 together with the blue one, as long as the red one might yet be in box 1. Things are different, however, if it is excluded that the red marble is in box 1. In that case, (2) holds. That is,

(4) *If the red marble is in box 2, then if the blue marble is in box 2 as well, the yellow marble is in box 1*

For those who do accept the last two statements there is a surprise in store. Using standard logical notation [1], we see that sentences (3) and (4) are respectively of the form

(3') $\sim$(*blue in 2* $\rightarrow$ *yellow in 1*)
(4') *red in 2* $\rightarrow$ (*blue in 2* $\rightarrow$ *yellow in 1*)

So we can apply the principle of Modus Tollens to (3') and (4') and conclude

(5') $\sim$*red in 2*

That is,

(5) *It is not the case that the red marble is in box 2*

But this conclusion will be a lot less acceptable than the premises (3) and (4) might have seemed.[2]

Is Modus Tollens playing up here, or should (3) or (4) be rejected after all? Any who have been confused by the above will feel obliged to choose sides.

Some may choose to stay with their initial intuitions about (3) and (4), regarding the example as evidence that Modus Tollens sometimes fails. But what kind of evidence is this? Why should intuitions be treated with this kind of respect, especially where others do not share them?

Choosing instead to do something about (3) and (4), and thus save the principle of Modus Tollens, would seem a lot safer: all [3)] existing theories of conditionals will support this. Even so, one would not have an easy time working things out. For although everyone is agreed that something must be done, there is no consensus to be found in the literature as to precisely what it should be. Some theories will advise you to reject (3), other ones to reject (4). Often, however, the response is more sophisticated, amounting to a denial that we are dealing with a proper instantiation of Modus Tollens here. Thus it can for example be argued that (3) is incorrectly formalized as (3'), as a negation of an implication; there is a hidden operator, and (3) really means something like 'it is not *necessarily* the case that if the blue marble is in box 2, the yellow one is in box 1'. Or it is argued that (4) is incorrectly translated into (4') as an embedded conditional sentence, that its meaning lies closer to something like 'if *both* the red *and* the blue marbles are in box 2, then the yellow one is in box 1'.

We will be returning to these matters at length in chapter 2, so we save the discussion up for there. Suffice it here to remark that it is not obvious how this partiality for Modus Tollens is to be justified. As above, by an appeal to intuition?

I.1.2. Logic as a descriptive science

The marbles puzzle has not been presented under the illusion that it would refute any theory of conditionals developed so far. Actually, it was presented under quite a different illusion, namely that it might help to put the aims and methods of one at present quite popular approach to the logical analysis of natural language in an unfavourable light.

Typical of this approach, in what follows to be known as Methodological Descriptivism, are

(I) a drive towards *descriptive* theories; that is attempts to distinguish systematically between the valid and invalid arguments within some class of arguments taken from a given natural language;

(II) the idea that such a theory can and should be tested by comparing what it has to say about the validity of the arguments it covers with the *intuitive* judgments of those who use the language concerned.

Key word in (I) is 'descriptive'. It is supposed that a logical theory should be descriptive - and that it *can* be so, the idea being that the logical researcher is faced with a class of arguments, some of which are, in fact, valid and some of which are, in fact, invalid. The object of the investigations is to discover where the division lies and, if possible, to find out why it lies there and not elsewhere.

'Descriptive' appears in the philosophy of logic in senses other than this, as a qualification not of logical theories but of the logical laws they sanctify. For instance, it sometimes crops up in discussions about the relation between logic and reality. Do the laws of logic tell us anything about the world? 'Yes, they do' is the descriptivist answer: the laws of logic add up to a compendium of the broadest traits of reality, and, in the last resort, owe their validity to their correctly describing these. One also frequently encounters the notion in discussions about the relation between logic and thought. There, a descriptivist is someone who tries to ground the validity of the laws of logic in the actual mental process of reasoning: 'logic is the physics of thought, or it is nothing'. [4)]

It is quite possible to be a descriptivist in the methodological sense while having nothing to do with either of these outmoded convictions about logical validity. Thus, the strictest of conventionalists can hold that the validity of the laws of logic is just a result of the linguistic conventions which govern the use of the logical constants, that other conventions about the use these words would yield other logical laws, and from this we may conclude that the laws of logic are not descriptive, neither of reality (not even of its basic structure [5)]), nor of mental process (not even of those occurring in the soundest mind) [6)]. Still, a conventionalist must be rated as a *methodological* descriptivist if we insist that logicians, in their analysis of a given language, should restrict themselves to describing the conventions which happen to bind the speakers of that language, and that they should refrain from anything like reforming those conventions.

So much for (I). (II) must be seen as a first attempt to answer the following question: assuming that it is reasonable to demand that a logical analysis should result in a descriptive theory, then how are we to find out whether any such purported description really conforms to the facts?

Admittedly, (II) calls for some elaboration. What are intuitive judgments, and why do they matter when a logical theory is put to the test, while non-intuitive judgments do not? Two different answers can be distinguished in the literature, a traditional and a modern one.

I.1.2.1. The rationalist tradition

To appreciate the traditional answer we must let ourselves be carried back to the time when nobody ever seriously reckoned with the possibility of there being various alternative logics.[7] At that time, no logician engaged himself in so many words with 'the logic of natural language'. One was, as it were, in search of the one and only logic (or believed that it had already been found in Aristotle's Syllogistics or Frege's Predicate Calculus), and that one logic was as a matter of course taken to be the logic of the natural languages.

Furthermore, it is important to realize that until about sixty years ago, semantics was of no more than marginal importance to the discipline of logic. Until then the core of a logical theory was given by a system of deduction, comprising axioms and rules of inference. These were couched in a more or less artificial language, of which often not even the syntax, let alone the semantics, was explicitly stated.

Now, a minimal requirement for a system of deduction - for *any* system of deduction, if at least it is presented as a *characterization* of some logic - is that it be reliable: every sentence that can be deduced from a given set of sentences must really follow from that set.

Traditionally, a system of deduction - every system of deduction ever presented as a characterization of the one and only logic - was held to be as reliable as its various axioms and rules of inference, and their reliability was thought to be *self evident*. Take, for example, the law of non-contradiction: *it is not the case that both* ϕ *and not* ϕ; wouldn't it be perverse to deny its validity? Or take the principle of Modus Ponens: from ϕ and *if* ϕ *then* ψ it follows that ψ; isn't it obvious that this is always the case, whatever sentences ϕ and ψ may be? Such principles simply cannot possess any better credentials than their self evidence; they cannot be rationally justified, for they are themselves the principles which any rational justification must presuppose.

Self evident axioms and rules of inference, if indeed they are self evident, will of course make any further analysis redundant. But how do we grasp the validity of these principles of reason if reason itself cannot help us out?

It is there that intuition comes in: if our normal intellectual faculties fail, something else must enable us to acquire knowledge of the principles of reason. (Actually, it must enable us to do so *a priori*, i.e. long before reason will ever feel the need to exploit them.) This idea of a special intellectual faculty - intuition - may sound rather *ad hoc*. But its significance for the history of philosophy can hardly be overestimated. Indeed, it is not too much to say that from Plato onwards, philosophers, in particular those standing in the rationalist tradition, have constantly been trying to make it into something more than an *ad hoc* solution.

For our purposes, it is not necessary to treat these matters in further detail. It suffices to note that the development of various non-equivalent systems of deduction has made it increasingly impopular to establish the reliability of a logical theory in the manner described above. Indeed, even if one sticks to the absolutist view that there is only one logic, one must at least admit that any traditional estimation of the time and the trouble necessary to find it would be overoptimistic. Self evidence has turned out to be an unworkable criterion.

In discussions of natural language various authors may, nevertheless, still regularly be heard advocating one or the other argument form as being intuitively valid. Even non-absolutists tend to do so as soon as the discussion is restricted to *the* logic of a given natural language. [8)] This, however, is by itself not enough to convict them of the views here described. For, although this usage of 'intuitively' is certainly rooted in the traditional usage, the word has in the meantime lost so much of its original impact that it might just as well be scrapped. Authors using 'intuitively' in such a context often mean no more than that the argument form in question seems reasonable enough to them, thereby not excluding the possibility that it might turn out to be invalid after all. In other words, in such a context 'intuitively' marks the introduction of an hypothesis rather than of an established truth.

Setting aside the descriptivist connotations, there is of course no fault to find with this usage of 'intuitive' just as long as one is aware that it would be begging the question to defend such an hypothesis by recourse to its intuitive validity if one is confronted with, for example, an argued counterexample. That would only lead to dogmatism. Or, whenever as in the case of conditionals, so many divergent views are held by so many authors, to an impasse.

I.1.2.2. The empirical approach

On to the second, more fashionable view of intuitive judgments. In practice it is not difficult to distinguish these from the first and traditional sort.

As we saw in the previous section, intuition was traditionally summoned in order to establish general logical principles. Consequently, traditional intuitive judgments - or at least those found worth recording - always say that some argument *form*, e.g. the Principle of Modus Tollens, is logically valid. Modern intuitive judgments, on the other hand, mainly serve the purpose of falsifying putative logical principles. They are judgments of *concrete* arguments (e.g. the particular instance of Modus Tollens given in section 1.1) which typically turn out to be intuitively 'absurd'.

Furthermore, 'intuitive' in the traditional sense goes with 'intuition', in the *singular*, sometimes preceded by 'our'. In the case of 'intuitive' in the modern sense, on the other hand, the *plural* 'intuition*s*' is employed, and mostly it is not our intuitions which are at issue but those of the native speakers of the language concerned. If an author speaks of 'our intuitions', a precautionary 'pre-theoretical' or 'untutored' will usually be inserted.

Indeed, 'intuitive' in the modern sense is directly opposed to 'theoretical'. Any judgment arrived at by straightforwardly applying some logical theory to a given argument is deemed non-intuitive. Of course, any such judgment is of no use whatsoever if it is the reliability of the theory applied or of any of its rivals which is at stake. Only the judgments of the theoretically unbiased can then be allowed to count - this in order to preserve the impartiality, one might almost say the objectivity, of the data against which the predictions of the theory concerned are to be tested.

Hence, intuitive judgments in the modern sense are judgments about the validity of concrete arguments made by theoretically unprejudiced speakers of the language concerned.

There must be more to intuitive judgments than this. How reliable are they? After all, it would seem that even the most impartial arbiter may be mistaken in her judgment. Can we not hope for more trustworthy data?

Commonly [9)] it is suggested that all competent speakers of a given language must, in virtue of their competence, be implicitly acquainted with its logical characteristics, and that it is this *subconscious knowledge* which surfaces in an intuitive judgment. Now clearly, any judgment betraying knowledge, even subconscious knowledge, must be correct. So, if this is what intuitive judgments do, then they are all true. Unfortunately, however, one cannot tell by just the form of a judgment whether it is a case of bona fide knowledge or merely one of belief. It is even impossible to decide whether ones own judgments were implicitly known to be true before they were explicitly believed to be so. Take the case of the marbles for example: do you know your own response to be the correct one, or do you merely believe this?

So this suggestion does not really help much. The gap between impartiality on the one hand and incorrigibility on the other appears unbridgeable, at least in practice. Or are we perhaps supposed to consult only speakers so competent that they are never mistaken? Then how should they be selected?

Some general criteria are of course readily available: very young children cannot be expected to have had the opportunity to develop their language skills, while others may for some reason not be able to. But it would seem that in order to discriminate any better than this, we will have to take recourse to *logical* criteria of some sort. Then, for example, speakers so incompetent that they can simultaneously believe the statements '*Jupiter is bigger than Mars*' and '*Mars is bigger than Jupiter*' could be excluded. As William Cooper (1978: 57), who has given a detailed exposition of the descriptivist view, puts it: 'If someone did claim to believe them both, one would have to challenge either his understanding of English, particularly his understanding of the full meaning of comparative construction, or else his intellectual capacity for applying his linguistic knowledge accurately in this particular situation.'

Well, take your pick. And then decide whether English speakers who simultaneously believe the statements (3) and (4) of section 1.1 should be treated in the same manner - or do they both understand the full meaning of 'if ... then' and 'not' and apply their linguistic knowledge accurately in this case?

No self respecting descriptivist will want to have anything to do with such selection procedures. Indeed, descriptivists will do their utmost to banish all logical bias, the more so as they expect their informants to do so as well. Or again, as Cooper (1978: 89) puts it, quite unaware that he might be contradicting his earlier remarks: 'In order to gain a more objective view of a language one must instead try to think like a Martian who has no idea what any of the human languages are like.'

In addition to the problem of ascertaining an informant's competence, there is also the problem of ascertaining his or hers impartiality. This is no less troublesome. For one thing, it is a fact that the intuitive judgments of those who have been in contact with logical theories are in many cases different from the judgments of those who have not. Somehow a training in logic affects ones powers of judgment: once speakers have been exposed in this way their intuitions appear to be corrupted for good, no matter how they will try to unburden their mind from its theoretical load. In any case, we can never be sure if and to what extent the affected persons have retained their original powers of judgment. So, for safety's sake all those acquainted with logical theories should be excluded from having any intuitive say. Moreover, it does sometimes happen that speakers confronted with some such quirk of language as the case of the marbles quite healthily and of themselves begin to theorize, in order to remove the confusion. Their considerations may be amateurish in comparison with those of the professional logicians, but they are no less infectious. So perhaps all those suspected of theoretical tendencies should be excluded as well. But where would all of this stop and who would be left over?

The most serious limitation of this modern approach lies, however, in the fact that professionals no less than amateurs usually invent logical theories precisely where their pre-theoretical intuitions desert them - and for that reason. Aristotle's Sea Fight Argument, the Sorites Paradox, the Paradoxes of Zeno, the Liar: all contain arguments which simply are not, intuitively and without further ado, valid or invalid. The situation may be less disastrous when dealing with, say, non-referring definite descriptions, or with conditionals, but even in these cases it would be gratuitous to suppose that we all intuitively know what we are doing. It is a riddle how our pre-theoretical intuitions, vague and dubious as they often are, could serve as a touchstone in such cases.

The only way out, it seems, would be to draw a distinction between clear intuitions and the less clear, with all of the problems of operationalizing which this would bring in its train. And even if these problems, which for the concepts of impartiality and competence would seem difficult enough, can be solved, it remains to be *shown* that the competent and impartial speakers of a language have sufficiently many of these 'clear' intuitions to ensure that there is only one logic covering them.

The descriptivists themselves were and are among the first to recognize these difficulties with the empirical basis of what they call Natural Logic, an empirical science which aims at discovering the logic(s) underlying natural language(s). Only they do not think that these difficulties are restricted to their discipline. As a rejoinder to the criticism that the intuitions of the native speakers of a language do not constitute a rock bottom empirical basis for testing logical theories they will be inclined to invoke philosophers of science such as Lakatos and Feyerabend who said that there aren't any rock bottoms anyway - not even for most physical theories. For example, the criticism brought forward in connection with the concept of competence might be considered tantamount to the remark that a logical theory is for its testing dependent on another theory - a theory of competence. And that is common enough in the field of empirical science. As Paul Feyerabend (1970: 204) puts it: 'It is hardly ever the case that theories are directly compared with 'the facts' or with 'the evidence'. What counts and what does not count as relevant evidence usually depends on the theory *as well* as on other subjects which may conveniently be called 'auxiliary sciences' ('touchstone theories' is Imre Lakatos's apt expression)'.

Maybe this analogy is instructive and maybe other analogies can be drawn between natural logic and established empirical disciplines in order to cover some of the other difficulties which I mentioned. But I do not think they can all be removed in this way. Take the problems I discussed in connection with the concept of impartiality. Certainly, it will not be difficult to find examples of measuring instruments which react like theoretically biased informants. To give an example, one cannot expect to falsify the statement that the volume of a fixed mass of fluid mercury at constant pressure is directly proportional to its temperature, in an experiment where the temperature is measured with the aid of a mercurial thermometer. Yet, measuring instruments which set their own standards, and even tend to change them, like informants tend to do who spontaneously start theorizing for themselves - those are unprecedented, I am afraid. Imagine a telescope lacking sufficient resolution, which when its resolution is inadequate, simply shows what it thinks might be there or should be there. If all telescopes worked that way, that would certainly cripple astronomy except if they did what they think best in a predictable manner (like the 'intelligent' television screens which increase contrast). But it is precisely this uniformity which cannot be expected in the case of self improving intuitions. People simply and as a matter of fact do not resolve the issues in the same way - otherwise there wouldn't be anything to argue about among the Natural Logicians.

I.1.3. A more pragmatic view

Suppose the logic of conditionals does exist. And suppose that we are presented with an attempt to describe it. Then the point made in the preceding pages is that we will have no way of seeing if the latter is an accurate representation of the former. At crucial points, the theoretically trained cannot be allowed to look, while the uninitiated will see only blurs.

In the pragmatic view, the logic of conditionals is not to be *recognized in* some logical theory; some theory is to achieve *recognition as* the logic of conditionals. On this account, a theory's ratification will have little to do with its accurately predicting intuitions and everything to do with its clarifying these. The better theory is the one so well motivated that people are prepared to allow it to *guide* their judgments whenever their intuitions leave them groping, and even to *correct* their judgments whenever their intuitions do not - not yet! - match the theory's predictions.

Pragmatists[10)] differ from descriptivists in not assuming that there already is a logic of conditionals, much less that the competent speakers, in virtue of their competence, are implicitly acquainted with it; their confusion when faced with the eccentricities of such sentences and the many disagreements, even among specialists, on the subject of conditionals are the pragmaticist's evidence for assuming that conditionals do not yet have any clear cut logic - in any case none which is accepted as such.

Now pragmatists and descriptivists differ in temperament as well. If the latter were by any chance to discover that conditionals do not yet have any logic - and surely, they must reckon with the possibility that the logic of conditionals might be simply unsettled at this point in the evolution of natural language - they would have to lay down tools. Being observers on principle, they can only wait and see if perhaps some new developments arise. Not so the pragmatists. Disregarding any advice against interfering with the natural development of language, they do approximately the following: they construct a theory which is intended as a *guide to using conditionals* and determine the logic conditionals would get if people were to use them the way this theory suggests.

Strangely enough, theories will not always carry the marks of their origin, whether descriptivist or pragmatist. The reason for this is that descriptivists are usually not satisfied with just *predicting* which arguments are valid and which are not; they want to *explain* things as well. In order to achieve this, they will often incorporate into their theory a *description of the way in which conditionals are used* - by the competent speakers, that is. But then, of course, descriptions of competent usage can easily be interpreted as guidelines for the less accomplished, and *vice versa*.

This, however, should not obscure the fact that pragmatists and descriptivists take quite different actions once their theories are ready. Descriptivists make it a point of duty to test the correctness of the theory by comparing its *predictions* with the intuitions of the competent speakers - a fool's errand, as we saw. To pragmatists the idea of testing a theory's correctness is quite foreign. They are hardly interested whether the speakers of the language concerned already use conditionals the way they would like them to, they want to know whether the speakers are willing to do so in future. And the only way to find that out is to see how they react to the theory's *explanations*.

Of course, pragmatists will not be able to force us into using conditionals according to their prescriptions any more than their descriptivist colleagues can compel us to remain doing so according to their own descriptions. Ultimately, we will have to sort that out for ourselves. The difference is that whereas descriptivists must be content tagging along behind changes in usage, pragmatists take it upon themselves to bring them about: they will try to ensure that following their advice would turn language into a more useful instrument of communication.

What about the judgments of the competent and theoretically unbiased speakers of the language - don't they matter anymore? *Of course they do*. And so do the judgments of the less competent and the theoretically biased speakers, though on the pragmatic account a speaker's judgment on the validity of a given argument, whether intuitive or not, is in many cases just the beginning of a test rather than the end. For example, if according to the theory in question a certain argument is logically valid whereas according to a certain informant it is not, then what matters is not so much the informant's judgment, but the considerations that have led to it. If there are no such considerations - i.e. if the informant can only take recourse to her intuitions - it may be helpful to explain how the validity comes about according to the theory in question. It might happen that after this explanation the informant changes her mind; she might even conclude that she was mistaken in her judgment. It might also occur that the informant still refuses to accept the validity of the argument. But at that stage of the discussion she will probably be able to be more explicit about her motives for doing so. The least one may expect her to do, then, is to point out what she does not like about the explanation so that one may get an indication as to why she does not want to use conditionals the way the theory suggests. It would be a lot more helpful, however, if apart from that she would offer an alternative theory and explain how her theory renders the argument invalid. Then it is the proponent of the original theory who might change his mind; he might be prepared to admit that the alternative theory suggests a better way of using conditionals than his own theory does. But it might also be that he sticks to his theory and adduces some new arguments in support of it in order to make his opponent as yet change her mind. And so on.

Admittedly, this picture lacks detail. It does not give any clue as to the kind of arguments that can play a role in the discussion. And perhaps it is too rosy a picture, too optimistic about the extent to which one or the other theory will emerge in the heat of the competition as a better theory than the others. Who is to say that the informant will allow the theoretical discussion to change her mind for her? Indeed, in this book we will on several occasions meet up with theories which are unsatisfactory because they render argument forms valid which we simply cannot get ourselves to support. The explanations which these theories give for this claim are simply not enough to bend our intuitions in their favour.

Still, the sketch given above does show how any disagreement in the validity of some concrete example naturally develops into a full-fledged theoretical discussion. And that is the point I wished to make. Theoretical discussion, something descriptivists at best condone as a marginal activity (also known as 'explaining away putative counterexamples'), is all in the day's work for a pragmaticist.

One final topic: does the pragmatic approach hold out better prospects for the problem of conditionals than the descriptivist approach? It may sound rhetorical, but this question must still be answered with some caution. Since it allows for a comparison of rival logical theories not only at the level of prediction but also at the level of explanation, the pragmatic approach enables us to evade the kind of problems that beset the descriptivist. So much, I hope, will have become clear. But new problems lie ahead. We are supposed to evaluate these theories according to the usefulness, fruitfulness, efficaciousness or what have you of the alternative ways of using conditionals which they prescribe. And at this point a good measure of scepticism may well be due. Do these notions provide any workable criteria? And if not, do they make much sense in this context?

I am afraid that the main reason why words like 'useful' and 'efficacious' slip out so easily when we are talking about ways of using 'if ... then' or other phrases, is that we are being carried away by a metaphor: words are like tools - and of course some tools are more useful than others, just as some ways of handling a given tool are more efficacious than others. In the case of real tools like hammers and saws this is fairly clear-cut, because it is obvious from the start what we want to use them *for*. Moreover, we can always perceive the results of applying them. Thus, we can literally demonstrate their utility. And if a certain way of handling such a tool is not very instrumental, we can always, so to speak, furnish material proof of this fact.

Now think of the results of applying a linguistic tool like 'if ... then'. Or try to explain what this tool can be used *for* - indeed, is 'if ... then' used *for* anything at all?

It is highly questionable whether the metaphor of tools puts us on the right track here. Still, the last question is meaningful, also when it is taken literally - and I remain very much inclined to answer it in the affirmative. There is more to a word than just *the way in which it is used*, there is also *the purpose for which it is used.*[11] In most cases the former is well-suited to the latter. But the very problem of conditionals is that however we use 'if ... then' (not in the least when we do so in an intuitive way), things never seem to work out quite as we want them to.

Unfortunately, the obvious next question - well then, what is the purpose of 'if ... then'; how should things work out? - is extremely difficult to answer in the abstract. Apparently, the only way to get to grips with it is by studying concrete proposals for using 'if ... then'. And even that may be too much said. The only thing that can be said for certain is that by comparing these proposals we can sometimes decide that one is more adequate to our needs than the other, without thereby being able to tell whether these needs were there from the beginning or whether they are newly arisen ones, aroused as it were by the comparison itself.

So the criterion of usefulness, even though it is ultimately decisive, is not very workable in practice. One cannot make up, in advance, a list of requirements that any useful way of using 'if ... then' should meet - no definitive list, that is. Besides, the relative usefulness of any proposed way of using 'if ... then' is certainly not the only thing that matters. For one thing, it might very well be that some of these proposals do not provide anything worth calling a way of using 'if ... then' (or any other expression for that matter) in the first place - not so at least in the eyes of someone whose own proposal is based on an altogether different philosophy of language. To put it otherwise: the question which theory suggests the most useful way of using 'if ... then' only becomes pertinent when we are comparing theories developed within the same theoretical framework. When the frameworks differ more general questions will arise, pertaining to the frameworks themselves rather than to the particular theories developed within them.

Since the problem of conditionals has prompted several alternative approaches to logic as a whole - not even the concept of validity has been kept out of harm's way - we shall more than once have an opportunity to discuss such general questions in the sequel. It will appear that not all of them can be decided on purely pragmatic grounds. Some will be rather metaphysical in character, other ones will concern epistemological matters. What I have called a theoretical framework contains among other things a description of the kind of circumstances (ontological and epistemological) in which the language users find themselves - in which, so to speak, they *cannot help* but find themselves. This kind of circumstances in its turn puts heavy constraints on the possible ways in which they might use 'if ... then'. Therefore, if we ever want to devise a way of using 'if ... then' that is of any use at all we'd better reckon with the circumstances as they really are.

The predictions of a logical theory are of minor importance, what matters is its explanatory force. For a logical theory has to be sold rather than tested. None of the comments made above should be allowed to obscure this fact. At best they tell that if we really want our theories to be bought, we must take care that they satisfy the actual needs of the language users. Not anything goes; there are all kinds of reasons why people might not be willing to buy a theory. Nevertheless, sometimes they do. Sometimes the way of speaking suggested by some logical theory is recognized by a large group of people as a useful way of speaking. Take classical first-order logic for example. Admittedly, the way of speaking underlying this logic does not serve all purposes equally well. Accordingly, the campaign to sell it as such, as was once the intention of Ideal Language Philosophy, has failed. Still, the way of speaking that goes with classical first-order logic has been recognized as most useful for *mathematical* purposes, in particular by those philosophers of mathematics who take a realist stand on the ontological status of mathematical objects. They *made* classical first-order logic the underlying logic of set theory. And now that set theory has become widely accepted as the basis of all mathematics, all students of mathematics are *taught* to express themselves as first-order logic says they should. In this book, too, especially in the more technical parts of it, classical first-order logic, together with the 'material' way of using 'if ... then' that goes with it, is adopted as the standard of reasoning for the metalanguage.[12)]

I.2. EXPLANATORY STRATEGIES

The leading theme of the previous chapter may not have made much of an impression on any who have themselves not experienced how deeply theoretical arguments can affect ones intuitions. Hopefully the following pages, which contain an extensive discussion of the case of the marbles will strengthen my point.

My principal concern, however, will be to introduce the various schools of thought in the field of conditional logic and to discuss the kinds of arguments which they employ. The marbles puzzle is not discussed for its own sake, but to this end.

I.2.1. Logical validity

I.2.1.1. The standard explanation

The usual explication of logical validity runs as follows:

(S1) *An argument is logically valid iff its premises cannot all be true without its conclusion being true as well*

By far the most theories of conditionals developed so far start from this explanation of logical validity; and nearly all of these are based on the additional presumption that non-truth equals falsity:

(S2) *No sentence is both true and false*
(S3) *Every sentence is either true or false*

Granted (S2) and (S3) the following truth-condition for negative sentences needs no further comment.

(S~) *A sentence of the form* $\sim\phi$ *is true iff* ϕ *itself is false*

Once you have come to accept all this, you hardly need a theory of conditionals to see that the Principle of Modus Tollens is logically valid. All you need is this sufficient condition for the *falsity* of a conditional sentence.

(SF) *Any sentence of the form* $\phi \rightarrow \psi$ *is false if its antecedent* ϕ *is true and its consequent* ψ *is false*

All theories based on (S1), (S2) and (S3) not only subscribe to (SF) but also to the following sufficient condition for the *truth* of a conditional sentence.

(ST) *Any sentence of the form* $\phi \rightarrow \psi$ *is true if its consequent* ψ *logically follows from its antecedent* ϕ

I.1. PROPOSITION. Given (S1), (S2), (S3), (S~), and (SF) the principle of Modus Tollens is logically valid.

PROOF. Suppose Modus Tollens is not valid in the sense of (S1). Then it should be possible for there to be three sentences of the form $\phi \rightarrow \psi$, $\sim\psi$ and $\sim\phi$ such that both the first and the second are true and the third is not true. If $\sim\phi$ is not true, then by (S~) ϕ is not false; and if ϕ is not false, then by (S3) ϕ is true. If $\sim\psi$ is true, then by (S ~) ψ is false. So ϕ is true and ψ is false, which with (SF) yields that $\phi \rightarrow \psi$ is false. Given (S2), this contradicts the requirement that $\phi \rightarrow \psi$ should be true. □

All theories based on (S1), (S2) and (S3) subscribe to (S~) and (SF). In other words, all of them sanctify Modus Tollens. Consequently their advocates will maintain that it is wrong to accept both the sentences (3) and (4) - or at any rate their formal counterparts (3') and (4') - occurring in the marbles puzzle. At least one of these premises must be rejected, but which one should it be? As I noted before, the consensus seems to dissolve here just as rapidly as it was reached. I shall present two theories, the one requiring us to accept (4') and not to accept (3'), the other one precisely the opposite.

The first theory tells us that 'if ... then' is just the so-called material implication: the condition laid down in (SF) is not only sufficient but also necessary for the falsity of conditional sentences. Using (S2) and (S3) this means:

($\supset$) *A sentence of the form $\phi \rightarrow \psi$ is true iff its antecedent ϕ is false or its consequent ψ is true*

This has certainly been the most disputed truth-condition in the history of logic ever since the Megarian Philo (fourth century B.C.) first suggested it - not only the most heavily criticised, but also the most ably defended. In recent introductory textbooks nobody ever pretends that ($\supset$) exhausts the meaning of 'if ... then'. At best it is pointed out that it is the only alternative left once one has chosen for (S2) and (S3), assuming one insists on speaking a *truth-functional* language, i.e. a language for which the following holds:

(T) *The truth value of any compound sentence is uniquely determined by the truth values of its constituent sentences*

Given (S2), (S3) and (T), the question as to which truth condition is best for conditional statements boils down to the question as to which of the values 'true' or 'false' the compound sentence $\phi \rightarrow \psi$ should be assigned in each of the following cases: (i) ϕ is true, ψ is true; (ii) ϕ is true, ψ is false; (iii) ϕ is false, ψ is true; (iv) ϕ is false, ψ is false. The answer is: (i) true; (ii) false; (iii) true; (iv) true. (Proof: consider the following conditional sentence *If Allen is over fifty years old, then he is over thirty years old.* We certainly want this sentence to come out 'true', whatever Allen's age may in fact be. Now suppose Allen is in fact sixty; then the antecedent is true, and so is the consequent (case (i)). If Allen is in fact forty, the antecedent is false and the consequent true (case (iii)). And if he turns out to be twenty, both the antecedent and the consequent are false. So there is at least one example of a conditional sentence that is true in case (i), true in case (iii) and true in case (iv). But if one conditional sentence is true in these cases, then by (T) all are. Hence, we must put the value 'true' in case (i), (iii) and (iv). That we must put the value 'false' in case (ii) is now obvious if only because otherwise every conditional sentence would turn out to be a logical truth.)

The logical theory of truth-functional languages is relatively simple. Therefore from a didactic point of view it is preferable to discuss the properties of those languages first in an introductory course. But does (T) have any merits over and above this one? Why would anyone *want* to speak a truth-functional language? The answer to this question is, I think, to be found not so much in (T) alone, but in the combination of (T) and

(A) *The truth value of any non-compound sentence is solely and entirely dependent on what is, in fact, the case*

(T) and (A) are two of the cornerstones of Wittgenstein's *Tractatus Logico-Philosophicus*. In a way, they just restate the old positivist principle that only facts can be a genuine source of knowledge: the truth value of any statement is wholly dependent on what is, in fact, the case. But (T) and (A) do more than just restate this principle, they also tell how one might live up to it: speak a truth-functional language, and things will work out exactly as you want them to. No wonder, then, that Wittgenstein's *Tractatus* found so much response among the members of the Vienna Circle.

We are ready now to apply this theory to the marbles puzzle. Applying (⊃) to (4'), we see that it is false just in case all three marbles are in box 2. But we already know that this cannot possibly be so. Only the situations depicted in the table can obtain, and whichever of these may happen to be the real situation, (4') will hold. Therefore, we can safely accept (4').

It is on the other hand not possible to say anything definite about the truth value of (3'). (⊃) and (S~) say that (3') is true iff both blue and yellow are in box 2. So (3') is true if the real situation is like the one depicted in III, but false if the real situation happens to be any of the others. Hence, it would be premature to say anything definite about (3') at this stage. We cannot accept it, but we cannot reject it either. This is not to say (3') has not got any definite truth value, only that we lack the information to decide which truth value.

The second theory I want to discuss would have us believe that 'if ... then' is a so-called strict implication. Roughly speaking, a conditional sentence is false according to this theory not only if it is *in fact* the case that its antecedent is true and its consequent false, but also if that might *possibly* be so. Just to get an idea of how this works. we turn to (3') again. Clearly, the blue marble might be in box 2 together with the yellow one (witness situation III). This means that the sentence *blue in 2* $\rightarrow$ *yellow in 1* evaluated as a strict conditional, turns out false. Consequently its negation (3') turns out true.

Before we can say what happens to (4') we must develop the rough idea given above more fully. Note first that the principle (T) of truth-functionality is abandoned by this theory: the truth value of a conditional is not uniquely determined by the truth values of its constituents - not by their *actual* truth values at least. In evaluating a conditional we must reckon not only with the truth values which its constituents happen, in fact, to have, but also with the truth values which these might possibly have. We are as it were to transfer ourselves to situations other than the actual one, and to evaluate the constituents there. If we are dealing with constituents that are themselves conditionals then evaluating the constituents in these other situations will involve transferring ourselves to yet 'other' other situations, and so on.

One way[13] of working this out is to say

($\Rightarrow$) *A sentence of the form* $\phi \rightarrow \psi$ *is true in a given possible situation s iff there is no possible situation s' such that* ϕ *is true in s' and* ψ *is false in s'*

Truth *simpliciter* is then to be understood as truth in the actual situation.

Now, as far as the case of the marbles is concerned, the relevant possible situations are the ones depicted in the table.[14] There is one situation, III, in which the blue and the yellow marble are both in box 2. So, by ($\Rightarrow$) we have that the sentence *(blue in 2 → yellow in 1)* is false in each of the situations I-VI. In some of these, notably I, II and VI, the red marble is in box 2. So there are situations in which the sentence *red in 2* is true while the sentence *(blue in 2 → yellow in 1)* is false. Given (⇒), this means that (4'), *red in 2 → (blue in 2 → yellow in 1)* is false in each of the situations I-VI. In particular we have that (4') is false in the actual situation, whichever situation that may be. Therefore, it is wrong to accept (4').

It is of interest to compare (4') with

(1') *(red in 2 ∧ blue in 2) → yellow in 1*

Given the following truth condition for conjunctions it can easily be verified that unlike (4'), (1') is true.

(S∧) *A sentence of the form $\phi \wedge \psi$ is true (in a given possible situation s) iff both ϕ and ψ are true (in s)*

Apparently the theory of strict implication distinguishes between sentences of the form $(\phi \wedge \psi) \rightarrow \chi$ and sentences of the form $\phi \rightarrow (\psi \rightarrow \chi)$. The theory of material implication does not, and - intuitively - that is a point in its favour. But the material implication has many problems of its own, as we will see in due course. So the question arises whether or not there are any other theories in the framework described by (S1), (S2), (S3), (S~), (S∧), (SF) and (ST) which render these two argument forms equivalent.

Its answer[15] is that there are not.

I.2. PROPOSITION. Within the framework given by (S1), (S2), (S3), (S~), (S∧), (SF) and (ST), the theory of material implication is the only theory which renders the argument form

(*) $(\phi \wedge \psi) \rightarrow \chi \;/\; \phi \rightarrow (\psi \rightarrow \chi)$

valid.

REMARK. Here and elsewhere I write '$\phi_1,\ldots,\phi_n$ / ψ' for the argument with the set of premises $\{\phi_1,\ldots,\phi_n\}$ and the conclusion ψ.

PROOF. We must show that the only truth condition for sentences of the form $\phi \rightarrow \psi$ left over is this:

$\phi \rightarrow \psi$ is true iff ϕ is false or ψ is true

The proof from left to right has already been given in the proof of proposition 1.

For the converse, we must show that

(i) if ϕ is false, then $\phi \rightarrow \psi$ is true;

(ii) if ψ is true, then $\phi \rightarrow \psi$ is true.

(i) Suppose ϕ is false. By (S~), it follows that $\sim\phi$ is true Note that given (S1), (S2), (S∧), and (S~), $(\sim\phi \wedge \phi)$ / χ is valid. In view of (ST) this means that $(\sim\phi \wedge \phi) \rightarrow \chi$ is true. Using (*) we see that $\sim\phi \rightarrow (\phi \rightarrow \chi)$ is true. Since $\sim\phi$ is true, it follows that $(\phi \rightarrow \psi)$ is not false, otherwise (SF) would not hold. So, by (S2) $\phi \rightarrow \psi$ is true.

(ii) Analogous. (Note that given (S1) and (S∧) $\psi \wedge \phi$ / ψ is valid.) □

I.2.1.2. Truth and evidence

In part II of this book we will return at some length to the standard concept of validity and the various theories of conditionals founded on it. In part III, a theory is developed on the basis of a different explication of validity. Compare (D1) with (S1):

(D1) *An argument is logically valid iff its premises cannot all be true on the basis of the available evidence without its conclusion being true on the basis of that evidence as well*

We obtain (D2), (D~) and (DF) in exactly the same way from (S2), (S~) and (SF) by substituting 'true on the basis of the available evidence' and 'false on the basis of the available evidence' for 'true' and 'false' in the original.

We cannot translate (S3) in this manner, since the result of doing so is unacceptable: not every sentence need to be decided, true or false, by the available evidence. For example, the data presently at your disposal neither allow you to conclude that the red marble is in box 2, nor that it is not.

(DF) offers only a sufficient condition for a sentence of the form $\phi \rightarrow \psi$ to be false on the basis of the available evidence. As a first approximation of a necessary and sufficient condition we have

(D→) *A sentence of the form* $\phi \rightarrow \psi$ *is false on the basis of the available evidence iff this evidence could develop into evidence on the basis of which* ϕ *is true and* ψ *is false. Otherwise,* $\phi \rightarrow \psi$ *is true on the basis of the available evidence.*

Let us apply these ideas to the marbles puzzle, beginning with (3'). Clearly we could at some later stage be less ignorant about the exact distribution of the marbles among the boxes. We might for example find out that the distribution in III of the table is the real one. There both the blue and the yellow marble are in box 2. This means that the evidence we have could develop into evidence on the basis of which the sentence *blue in 2* is true and the sentence *yellow in 1* is false. So, by (D→), *blue in 1* → *yellow in 1* is false on the basis of the evidence presently available and with (D~) this gives that its negation ~(*blue in 2* → *yellow in 1*) is true.

As far as (4') is concerned, suppose that our present information about the distribution of the marbles was to grow in such a way that we knew its antecedent, *red in 2*, to be true. *Then* (4')'s consequent *blue in 2* → *yellow in 1* would also be true on the basis of the available evidence. This is easily seen as follows. If the information could grow some more in such a way that we were to learn that the blue marble is in box 2 and the yellow is not in box 1, we would at that stage have evidence on the basis of which all three sentences '*red in 2*', '*blue in 2*', '*yellow in 2*' were true. But our initial information that there is at least one marble in each box excludes this possibility.

According to this theory, then, both (3') and (4') are quite acceptable, while Modus Tollens fails. It does not follow from (3') and (4') that the red marble is not in box 2; we only get that it *may* not be there.

Still, this theory does not leave us very much room to tinker with Modus Tollens. (D1), (D2), (D$\sim$) and (D$\rightarrow$) preclude three sentences of the form $\sim\psi$, $\phi \rightarrow \psi$ and ϕ all being *true* on the basis of the same evidence. So, what happens if we are finished with the extra information that the red marble is in box 2? Well, hopefully something will happen to the truth of (3') or (4'). It does, as can easily be verified. On the basis of such new information, (3') will be false. What this means is that truth or falsity on the basis of the available evidence need not be invariable under growth of that evidence. Various sorts of sentences do possess this sort of stability, but conditional sentences typically do not. A conditional $\phi \rightarrow \psi$ can be false on the available evidence simply because it is *not yet* possible to rule out the possibility that ϕ will turn out true without ψ turning out true too: the available evidence is just too scanty. Adding information which does rule this possibility out will then switch the truth value of the conditional.

It is worth noticing that this approach in a sense lies somewhere between conditionals as material implications and conditionals as strict implications. In evaluating a material implication, one is only interested in what holds in reality - one behaves, so to speak, as if the evidence is complete. In evaluating a strict implication, one always takes *all* possibilities into account - as if, so to speak, one never learns that some of these are, in fact, excluded. But according to this approach, one is interested only in those possibilities which are left open by the evidence which one happens to have. This, of course, makes conditionals highly context dependent. In a context where no specific evidence is available, they are like strict implications. In a context where the evidence is complete, they are like material implications. In most contexts they are neither.

I.2.1.3. Probability semantics

The theory developed in part III is not the first theory of conditionals based on a non-standard explication of logical validity. Indeed, there is something about conditionals which seems to invite such a manoeuvre.

Ernest Adams (1966, 1975) has also proposed a modification of the classical standard of logical validity. His idea is that some conclusion logically follows from a set of premises not if its *truth* is guaranteed by theirs, as the classical standard would have it, but rather its *probability*. An argument is said to be valid, on Adam's view, if it is possible to bring the probability of its conclusion arbitrarily close to one by raising that of its premises above some suitable value. More precisely,

(P1) *An argument* Δ/ϕ *is valid iff for every* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that for any probability assignment* P *if* $P(\psi) > 1-\delta$ *for each* $\psi \in \Delta$, *then* $P(\phi) > 1-\varepsilon$.

Consider the set L_0 of sentences that can be formed out of the atomic *red in 1*, *blue in 1*, *yellow in 1*, *red in 2*, *blue in 2*, and *yellow in 2* by conjunction and negation only. The next definition says what a probability assignment to the sentences of L_0 is.

(P2) *A probability assignment for* L_0 *is a function that assigns a value* $P(\phi)$, $0 \leq P(\phi) \leq 1$, *to every* $\phi \in L_0$, *while furthermore*

(*i*) $P(\phi) = P(\phi \wedge \phi)$

(*ii*) $P(\phi \wedge \psi) = P(\psi \wedge \phi)$

(*iii*) $P(\phi \wedge (\psi \wedge \chi)) = P((\phi \wedge \psi) \wedge \chi)$

(*iv*) $P(\phi) + P(\sim\phi) = 1$

(*v*) $P(\phi) = P(\phi \wedge \psi) + P(\phi \wedge \sim\psi)$

Think of the probability $P(\phi)$ of a sentence ϕ as an agent's *degree of belief* in ϕ - something that could be measured by betting odds: if you are willing to bet at odds of 1:5 - but at no higher odds say 1:1 - for the proposition that both the red and the blue marble are in box 2, then your degree of belief $P(\textit{red in 2} \wedge \textit{blue in 2})$ = 1/1+5 = 1/6 - rather than 1/1+1 = 1/2.

(P2) imposes heavy constraints on the way an agent can distribute his degrees of belief among the various sentences of L. Actually, it is not meant as a description of how a particular agent will do so, but of how a *rational* agent would. Some distributions of belief are not rational in that they allow a Dutch book to be made against anyone whose state of beliefs conforms to them. (A Dutch book is a set of bets which the holder of the beliefs must accept (given his belief distribution) but which will certainly result in a net loss to him in the long run.) It can be shown that a necessary and sufficient condition of not being in a position to have a book made against you is that your degrees of belief in the sentences of L_0 satisfy the axioms laid down in (P2), which are just a version of the axioms of the probability calculus.

Example: Assuming that you find each of the six distributions of the marbles among the two boxes equally probable (and assuming you do not consider any other distribution possible), you should get

$P(\textit{red in 1}) = 1/2$

$P(\sim(\textit{red in 2} \wedge \textit{blue in 2})) = 5/6$

Now, what is the logic generated by (P1) and (P2)?

I.3. PROPOSITION. Let Δ/ϕ be an argument of L_0.
Δ/ϕ is valid according to (P1) and (P2) if and only if Δ/ϕ is valid according to (S1), (S2), (S3), (S~) and (S∧).

PROOF. For a proof of this proposition the reader is referred to Adams (1975: 57-58). □

In other words, as far as L_0 is concerned, the probabilistic account does not bring up anything new. Things change, however, if we turn to sentences containing 'if'. The basis for the probabilistic treatment of the conditional sentences consists, not surprisingly, in the idea of conditional probability.

(P3) *Let* P *be a probability assignment to the sentences of* L_0 *and* ϕ *a sentence of* L_0 *with* $P(\phi) \neq 0$. *We define* P_ϕ *as the function assigning to every sentence* ψ *of* L_0 *the value*

$$P_\phi(\psi) = \frac{P(\phi \wedge \psi)}{P(\phi)}$$

I.4. PROPOSITION. Let P and P_ϕ be as above.
P_ϕ is a probability assignment in the sense of (P2).

PROOF. Left to the reader. □

$P_\phi(\psi)$ is called the conditional probability of ψ, given ϕ. Intuitively what P_ϕ is supposed to describe is the *change* of belief that results once ϕ is known to be true. Again it can be proved that agents who do not change their beliefs by 'conditionalizing' on this new information can always have a Dutch book made against them.

Example: Assuming, once again, that your degree of belief in each of the six possible distributions of the marbles among the two boxes is 1/6, you should get

$$P_{\textit{blue in 2}}(\textit{yellow in 1}) = 2/3$$

$$P_{\textit{red in 2}_{\textit{blue in 2}}}(\textit{yellow in 1}) = 1$$

$$P_{\textit{red in 2} \wedge \textit{blue in 2}}(\textit{yellow in 1}) = 1$$

Notice that $P_{\phi_\psi}(\chi) = P_{\phi\wedge\psi}(\chi)$ for any ϕ, ψ, χ such that $P(\phi \wedge \psi) \neq 0$.

Now let us see how we must adapt (P2) if L_0 is extended to L, the set of sentences that can be built from atoms *red in 1*, *red in 2* etc. using conjunction, negation and implication (and nothing else).

The basic idea is to equate the absolute probability of the conditional $\phi \rightarrow \psi$ with the conditional probability of ψ, given ϕ.

(P4) *If* $P(\phi) \neq 0$, *then* $P(\phi \rightarrow \psi) = P_\phi(\psi)$

If $P(\phi) = 0$ we leave $P(\phi \rightarrow \psi)$ undefined.[16)] (This means of course that we will have to reformulate the axioms given in (P2) in such a way that they do not apply to sentences whose probabilities are undefined.)

As an example of this idea working at its best, consider the Hypothetical Syllogism

$\psi \rightarrow \chi,\ \phi \rightarrow \psi\ /\ \phi \rightarrow \chi$

The relevant probabilities are then the following

$$P(\psi \rightarrow \chi) = \frac{P(\psi \wedge \chi)}{P(\psi)},$$

$$P(\phi \rightarrow \psi) = \frac{P(\phi \wedge \psi)}{P(\phi)}, \text{ and}$$

$$P(\phi \rightarrow \chi) = \frac{P(\phi \wedge \chi)}{P(\phi)}$$

We can choose mutually exclusive ϕ and χ, and a sentence ψ which is compatible with both of these. Then $P(\phi \rightarrow \chi)$ will be zero. But the probabilities of $P(\psi \rightarrow \chi)$ and $P(\phi \rightarrow \psi)$ can increase without this having any effect on $P(\phi \rightarrow \chi)$. So the Hypothetical Syllogism is not in general valid in the sense of (P1) which is surprising but just as well. No one would accept that from

If Jones wins the election then Smith will retire to private life

and

If Smith dies before the election then Jones will win it

it follows that

If Smith dies before the election he will retire to private life

(P4), however, cannot be all there is to it. For one thing, it does not allow us to extend the result mentioned in proposition 4 to the language L. That is, we cannot be sure that we still get an (extended) probability assignment after conditionalizing on an (extended) probability assignment.

It seems obvious that the way to achieve this would be to add the following requirement:

(P5) *If* $P(\phi \wedge \psi) \neq 0$ *then* $P_{\phi}(\psi \rightarrow \chi) = P_{\phi\wedge\psi}(\chi)$

Unfortunately, this does not work. As David Lewis (1976) showed, this way you end up with probability functions which are at most four valued.

I.5. PROPOSITION. Let P be an (extended) probability assignment. Suppose there are sentences ϕ and ψ such that both $P(\phi \wedge \psi) > 0$ and $P(\phi \wedge \sim\psi) > 0$. Then $P_{\phi}(\psi) = P(\psi)$.

PROOF. Note first that $P(\phi) > 0$, $P(\psi) > 0$, and $P(\sim\psi) > 0$. Therefore the following makes sense.

By (P5) and (P3) we have

$$P_{\psi}(\phi \rightarrow \psi) = P_{\phi\wedge\psi}(\psi) = \frac{P(\phi \wedge \psi \wedge \psi)}{P(\phi \wedge \psi)} = 1$$

$$P_{\sim\psi}(\phi \rightarrow \psi) = P_{\phi\wedge\sim\psi}(\psi) = \frac{P(\phi \wedge \sim\psi \wedge \psi)}{P(\phi \wedge \sim\psi)} = 0$$

Furthermore, by (v) of (P2), (P3) and (P4),

$$\begin{aligned} P_{\phi}(\psi) &= P(\phi \rightarrow \psi) \\ &= P_{\psi}(\phi \rightarrow \psi) \cdot P(\psi) + P_{\sim\psi}(\phi \rightarrow \psi) \cdot P(\sim\psi) \\ &= 1 \cdot P(\psi) + 0 \cdot P(\sim\psi) \\ &= P(\psi) \end{aligned}$$

□

I happen to believe that P(*red in 1* $\wedge$ *blue in 1*) = 1/6, and that P(*red in 1* $\wedge\sim$*blue in 1*) = 1/3. Furthermore I would say that $P_{\textit{red in 1}}$(*blue in 1*) = 1/3, and that P(*blue in 1*) = 1/2. But the above forbids this. Lewis (1976) puts it more generally: '... if we take three pairwise incompatible sentences ϕ, ψ and χ [I replace Lewis's notation by mine here, F.V.] such that $P(\phi)$, $P(\psi)$ and $P(\chi)$ are all positive and if we take Θ as the disjunction $\phi \vee \psi$, then $P(\Theta \wedge \psi)$ and $P(\Theta \wedge \sim\psi)$ are positive but $P_{\Theta}(\psi)$ and $P(\psi)$ are unequal. So there are no such three sentences. Further, P has at most four different values. Else there would be two different values of P, x and y, strictly intermediate between 0 and 1 such that $x + y \neq 1$. But then if $P(\phi) = x$ and $P(\psi) = y$ it follows that at least three of $P(\phi \wedge \psi)$, $P(\sim\phi \wedge \psi)$, $P(\phi \wedge \sim\psi)$, and $P(\sim\phi \wedge \sim\psi)$ are positive, which we have seen impossible.

The reaction of Adams to this perplexing triviality result is to shrink the domain of application of his theory such that the above argument cannot be set up. Starting from the idea that an assertion of a conditional is a conditional assertion and that as such conditionals lack the truth-values of ordinary assertions, he argues that condition (P4) only holds for conditionals $\phi \rightarrow \psi$ whose antecedent ϕ and consequent ψ do not contain other conditionals, and that it is wrong to ask for a generalization to other cases. Only unconditional consequents can be asserted conditionally and that only on non-conditional conditions. He even denies that one can attach probabilities to conjunctions, negations and other truth-functional compounds of conditionals.[17)]

As a consequence Adams' theory cannot help us solve the puzzle of the marbles. Both (3') and (4') fall outside the scope of his theory.

Lewis' triviality result does not only pose a problem for Adams, but for everyone who wants to attach probabilities to conditionals. In what follows I will not have much to say on this problem.[18)] Suffice it to say that from the perspective of data semantics it seems misguided to try attaching probabilities to conditionals. Roughly the argument is this: bets can only be laid on sentences that are stable in the sense that once they have turned out to be true/false on the basis of the available evidence, they remain so. In the preceding section we saw that conditionals do not have this property. Now, at which point will it be decided who has won?

I.2.1.4. Relevance logic

Another and quite different criticism of the standard notion of logical validity is to be gleaned from the work of the relevance logicians. They believe that for an argument to be valid, it is not sufficient that the truth of the premises be transferred to the conclusion. The premises of the argument must in addition be *relevant* to its conclusion. There is something in this. It is at least misleading to conclude from the irrelevant coincidence of it raining in Ipanema that the red marble either is or is not to be found in box 1. As Anderson and Belnap (1975: 14) put it:

'Saying that ψ is true on the irrelevant assumption that ϕ is not to deduce ψ from ϕ, nor to establish that ϕ implies ψ in any sensible sense of implies. Of course we *can* say *Assume that snow is puce. Seven is a prime number.* But if we say *Assume that snow is puce. It follows that* (or *consequently*, or *therefore*, or *it may validly be inferred that*) *seven is a prime number*, then we have spoken falsily.'

Under this banner they embarked on the ambitious programme of analyzing the relation of entailment in such a way as to circumvent these and other 'fallacies of relevance'.

The explication of logical validity developed in part III is not going to satisfy the relevance logicians any more than the classical one, and for the same reasons. It does not take any account of the relevance of the premises of an argument to its conclusion in assessing its validity. And this is not the only fault which they will find.

We turn to the marbles for an example. The argument

The red marble is in box 1

If it is raining in Ipanema, then the red marble is in box 1

is valid according to standards set in part III. The problem which relevance logic would have with this is not that its premise is irrelevant to its conclusion, but rather that the antecedent of the latter is irrelevant to its consequent. (As a matter of fact, relevance logicians scarcely distinguish these two levels.) As a result, the conclusion will be deemed false and the argument form will be deemed invalid.

Everyone would agree that to argue irrelevantly is to argue badly, as it is to argue from false premises, in a roundabout way, or to the wrong conclusion. And the claim that the antecedents of conditionals should be relevant to their consequents also has something to say for it. In any case, as anyone who has ever taught truth tables knows, this idea appeals to a wider group than just the relevance logicians (compare this with section III.2), and it is to their credit that they have insisted that these matters should not be forgotten.

It seems to me, however, that the difference between finding an argument invalid because the premises are irrelevant to the conclusion, and finding it valid though ineffective for the same reasons, is largely a verbal one. Besides this, from a methodological point of view, it is dubious whether there are any advantages in lumping together these various ways in which arguments can be improper. The relevance logicians run the risk of turning logical validity into a clumsy thing. The difficulties they have in providing their largely proof-theoretic theories with a proper semantics may be regarded as a symptom of this. The semantic theories which have thus far been put forward tend to lack the explanatory power which is to be expected from theories which purport to say what relevance means. They are in a sense *merely* formal, and are extremely difficult to apply in analyzing the sorts of things which we are interested in here. This applies not only to the larger part of the work done in this tradition, which is primarily concerned with the abstract notions of relevance and entailment, but also to the work done by Barker (1969) and Bacon (1971), which does focus attention on conditional sentences derived from natural language. In trying to apply their ideas to the marbles puzzle, for example, I have not been able to decide which of (3') and (4') they would recommend rejecting. (They must reject at least one of the two, since Modus Tollens is valid in relevance logic.)[19]

I.2.2. Pragmatic correctness

One cannot assert any sentence at any time; statements can be conversationally out of place even though they are true, highly probable, or true on the basis of the available evidence. Having made this trifling observation - after all 7 + 5 = 12 - we might ask for the criteria by which it can be determined whether or not a statement is conversationally correct. Following Grice (1975) we might try to find these criteria in some *maxims of conversation* which the participants in a conversation should observe in order that their conversation be as productive as possible.[20] Then, having found these criteria, we might carry on and try to map out the circumstances under which various kinds of statements can properly be made. In doing so we would discover that statements when used correctly convey much more information than just their *logical content*. In addition to this there are also the *pragmatic implicatures*. For example, we might find that an indicative conditional *if it is the case that* ϕ *then it is the case that* ψ usually 'implicates' both *it may be the case that* ϕ and *it may be the case that not* ϕ - usually, but not always:

implicatures are *cancellable*. When one asserts

if she is under twenty, then I'll eat my hat

one is not intending to suggest that she may be under twenty at all. And the statement

she is on the wrong side of thirty, if she is a day

does not implicate that she may be less than a day old.

We would also discover that sometimes it is very hard to distinguish between logical consequences and pragmatic implicatures; especially if we are not yet completely sure *which* logic we are dealing with. Is $\phi \rightarrow \psi$ a logical or a pragmatic consequence of $\sim\phi \vee \psi$? If you think that $\rightarrow$ behaves as a material implication, then undoubtedly you will say that it is the first. But if like Adams you believe that $\rightarrow$ behaves as a conditional probability, in which case you'll find the inference pattern $\sim\phi \vee \psi \ / \ \phi \rightarrow \psi$ logically invalid, you will agree that it is the latter. As Adams (1966: 285) puts it:

'What the present theory shows is that inferring *if* ϕ *then* ψ from *either not* ϕ *or* ψ is not *always* reasonable, but that the only situation under with *either not* ϕ *or* ψ has a high probability, but *if* ϕ *then* ψ has a low one is the situation in which *not* ϕ has a high probability. Assuming this, we have an immediate explanation of why we are ordinarily willing to infer *if* ϕ *then* ψ from *either not* ϕ *or* ψ: the reason is that people do not ordinarily assert a disjunction when they are in a position to assert one of its members outright. (In fact, it is misleading to do so, and therefore doing it probably runs against strong conventions for the proper use of language.) Thus, if one heard it said that *either the game will not be played tomorrow, or the Dodgers will win* he would be well justified in inferring *if the game is played tomorrow, then the Dodgers will win*, and what would justify the inference would be the knowledge that the person asserting *either the game will not be played or the Dodgers will win* did not do so simply on the grounds of the information he had to the effect that the game would not be played'.

Here we see pragmatic considerations being invoked to explain why a certain logically invalid inference pattern has so many intuitively sound instances. And below we see David Lewis (1976: 137) invoking exactly the same pragmatic considerations in order to explain why a certain logically valid inference pattern has so many intuitive counterexamples It concerns the scheme $\sim\phi$ / $\phi \rightarrow \psi$.

'The speaker ought not to assert the conditional if he believes it to be true predominantly because he believes its antecedent to be false, so that its probability of truth consists mostly of its probability of vacuous truth. It is pointless to do so. And if it is pointless, then also it is worse than pointless: it is misleading. The hearer, trusting the speaker not to assert pointlessly, will assume that he has not done so. The hearer may then wrongly infer that the speaker had additional reason to believe that the conditional is true, over and above his disbelief in the antecedent.'

As these examples illustrate, Gricean arguments have become standard repertoire in defending logical theories. The idea of a pragmatic theory complementing a semantic theory has become quite familiar. Now, everyone is in agreement that semantics and pragmatics should cooperate in this way, but there is a lot less agreement as to the distribution of labour among the two. The problem is that it is hard to say where semantics stops and pragmatics takes over. What the one author classifies as a clearcut counterexample to a putative logical principle is for the other merely an innocent pragmatic exception to an otherwise faultless logical rule.

Are there any general criteria which can be used to decide who is right and who is wrong, to distinguish the domains of semantics and pragmatics? I have come to believe that only global criteria would be of any use in this, that is doesn't make much sense to take any particular inference rule and to fight it out over that one isolated example. Instead, what matters is the way a combination of a logic and its complementary pragmatic theory performs in general. The best combination will be something like the combination which best *explains* the plausibility of as many plausible sounding inferences as possible, and best *explains* the implausibility of the rest. And the best dividing line between semantics and pragmatics will be the line drawn by whichever combination this is.

Of course this is both a simplified and rosy view of the matter. In practice, we do not have dividing lines, but gaps. Take for example classical logic. There are plenty of clearcut intuitive counterexamples to the classically valid scheme $\sim(\varphi \rightarrow \psi)$ / φ. For instance,

It is not the case that if the peace treaty is signed, war will be prevented

---

The peace treaty will be signed

As yet, however, no adequate pragmatic explanation in terms of maxims being overruled by anyone arguing in this manner has been provided. As a second example, take the classically valid scheme $(\varphi \wedge \psi) \rightarrow \chi$ / $(\varphi \rightarrow \chi) \vee (\psi \rightarrow \chi)$ and its following instantiation:[21)]

If both the mainswitch and the auxiliary switch are on, the motor is on

---

If the mainswitch is on the motor is on, or, if the auxiliary switch is on, the motor is on

It has proved extremely difficult to give a pragmatic explanation of what is going wrong here.

Still, I think that it should be possible to draw a neat line between semantics and pragmatics, and not leave gaps like these. Even better, I think that the semantic theory developed in part III of this dissertation draws such a line, and that this line is drawn exactly as Grice's theory of conversation prescribes: Every counterexample to an argument form dubbed logically valid is to be explained as a product of a violation of the conversational maxims. (Note that Lewis is giving such an explanation). And every argument form dubbed merely pragmatically correct, must have instantiations which show that the conclusion is sometimes cancellable. (Note that Adams does not give such an example.)

I.2.3. Logical form

The following quotation is taken from Geach (1976: 89).

'Roughly speaking, hypotheticals are sentences joined together with an 'if'. We don't count, however, sentences like *I paid you back that fiver, if you remember; There's whisky in the decanter if you want a drink;* for here the speaker is committed to asserting outright - not *if* something else is so - *I paid you back that fiver* or *There's whisky in the decanter.* Nor do we count sentences where 'if' means 'whether': *I doubt if he'll come* (quite good English, whatever nagging schoolmasters say). Nor do we count cases where 'if' has to be paraphrased with 'and': *If you say that, he may hit you = Possibly(you'll say that* and *he'll hit you); If it rains it sometimes thunders = Sometimes(it rains* and *it thunders).'*

Ever since 1905 and the publication of Russell's 'On denoting', the sort of distinction which Geach is implicitly making here between grammatical form on the one hand and logical form on the other has been quite familiar. For a while, during the heady youth of analytical philosophy, it even looked as if elucidating the logical form behind various misleading kinds of expressions would turn out to be the proper task of philosophy. It was a time in which one was largely interested in weeding out philosophy, in showing that large parts of traditional philosophy are in fact meaningless, a time in which Carnap could hope to show that Heidegger's work is nonsense just because it cannot be properly translated into standard predicate logical form.

These days we know that there is a lot more besides Heidegger which could not be translated into predicate logic, and not all of it is nonsense. Be this as it may, the notion of logical form is still very much with us, albeit in a modified role.

We have already mentioned the way a theory of pragmatics can and should complement a semantic theory, such that the two together cover as much of the whole field of what might be called intuitively sound argument forms as possible. In practice this does of course not always work out that well; an example of this is to be found in Adams' treatment of the Hypothetical Syllogism: $\phi \rightarrow \psi$, $\psi \rightarrow \chi$ / $\phi \rightarrow \chi$. According to his logical theory it is not valid, while he cannot think of a pragmatical reason why arguments of this form in many cases seem acceptable. The way he then tacitly invokes the notion of logical form in order to reformulate the premises so that the argument becomes valid is typical of the role which it is all too often given in the literature on conditionals. (See Adams,1975: 22).

"... we suggest that the 'hypothesis' of the first premise (the antecedent of the conditional) is tacitly 'presupposed' in the second, ... we will not attempt a rigorous justification of the foregoing intuitively plausible suggestion, but we will now see that if the suggestion is correct it would explain why apparent Hypothetical Syllogism inferences are rational, ... making the tacit presupposition of the second premise of a real life like Hypothetical Syllogism explicit, transforms it into an instance of the Restricted Hypothetical Syllogism pattern $\phi \rightarrow \psi, (\psi \wedge \phi) \rightarrow \chi / \phi \rightarrow \chi$ which is universally probabilistically sound ..."

At its worst this is a strategy which cannot but result in a proliferation of epicycles, and Adams is by no means the only one. Cooper (1978:199) is embarrassed that his theory deems $\sim(\phi \rightarrow \psi)$ and $(\phi \rightarrow \sim\psi)$ equivalent:

"It seems reasonable to challenge

It is not the case that if Jones' car is gone he is out

If Jones' car is gone he is not out

It is unclear (to me) just what is going on in examples like these. Perhaps *It is not the case that* when followed by a conditional statement is sometimes understood to mean *It is not necessarily the case that*. Or perhaps negations of whole conditional statements, being rare in English, have an interpretation which is idiosyncratic and simply unsettled at this point in the evolution of the language."

Recall that a similar jump from a negated conditional to an underlying logical form involving *necessary* was suggested in the first discussion of the marbles puzzle. It is a suggestion to which I am quite partial. As a matter of fact I wonder whether there are any negated English conditionals *without* this so-called hidden operator, but more about that presently. What I object to here is the ad-hoc manner in which such a possibility is introduced in order to save faces. If there are hidden operators in the negations of conditionals then the *theory* should be saying this. We shouldn't be forced to say it for the theory just to keep it sounding plausible.

A third example of the notion of logical form being used to patch up logical theories is to be found in the extensive literature on the argument form

(*) $(\varphi \vee \psi) \rightarrow \chi \ / \ \varphi \rightarrow \chi$,

which is invalid in most treatments of counterfactuals. This is usually thought unsatisfactory, but it turns out that you can render this argument form valid only at a fairly heavy price. If your theory says that logically equivalent sentences are interchangeable and that any sentence φ is logically equivalent to $(\varphi \wedge \psi) \vee (\varphi \wedge \sim\psi)$, then the validity of (*) implies the validity of

(**) $\varphi \rightarrow \chi \ / \ (\varphi \wedge \psi) \rightarrow \chi$,

which in the context of counterfactuals is a lot less attractive. (Write $(\varphi \wedge \psi) \vee (\varphi \wedge \sim\psi)$ for its equivalent φ in $\varphi \rightarrow \chi$ and then apply the principle in question once).

The reactions to this have been both many and varied. To be brief: Nute (1978, 1980) started by giving up the replaceability of logical equivalents, but later (1980 a) changed his position and defended (*) as pragmatically correct rather than logically valid. McKay and van Inwagen (1977) invented some counterexamples to (*), but most people think these are just pragmatically incorrect instances. Warmbröd (1981) accepts (**) and he finds its counterexamples pragmatically incorrect. But perhaps the most common response (see Loewer (1978), Ellis (1979)) is just to deny that natural language counterfactuals with disjunctive antecedents can properly be formalized as $(\varphi \vee \psi) \rightarrow \chi$. Such a counterfactual is held to have the logical form $(\varphi \rightarrow \chi) \wedge (\psi \rightarrow \chi)$, and then the rest is easy enough.

In the following I will try my utmost not to use arguments like the above. Only if it is clear that dealing with an *ambiguous* sentence like for example

it is not the case that the red marble is in box 1 if the yellow marble is in box 2

will I allow different logical forms: ~(yellow in 2 → red in 1) and (yellow in 2 → ~red in 1). But I cannot see any ambiguity in Geach's example: *If you say that, he may hit you,* therefore I will formalize it as a sentence of the form $(\varphi \rightarrow$ *may* $\psi)$ - that is, as closely to the surface structure as possible. For the same reason I will formalize English sentences which run like '*it is not the case that if ..., then ...*' as ~(... → ...). If you were to say these English sentences say that it is not *necessarily* the case that if ... then ...., then I would agree. But if you were to conclude from that that they should be formalized as ~□ (... → ...), then at least you should be able to come up with an example of a negated conditional sentence which does *not* say that it is not necessarily so that if ... then ... .

The reason why I follow this strategy is because I think it is the most sensible one to follow in studies like these, where hardly any attention is paid to syntactic questions. It is not because I think that there is no need for a 'level of logical form' in syntax, or that such a level is only needed for desambiguation, (as it is in Montague Grammar[22)]) I am ready to admit that syntax is just as important to semantics as pragmatics is, and that ultimately the question is which combined theory of syntax + semantics + pragmatics offers the best explanations. It might be that such a *theory* will enable us to use syntactic arguments in our explanations. It might even be that the notion of logical form will then be of key importance.[23] In the absence of such a theory, however, it is best to avoid ad-hoc explanations.

NOTES TO PART I

1. ~ is short for *not*, ∧ is short for *and*, ∨ is short for *or*, and → is short for *if ... then ...*

2. This example is a slight variant of Lewis Carroll's Barbershop Paradox, which first appeared in Carroll (1894).

3. Modus Tollens does fail in the theory of conditionals put forward in Cooper (1978). For this particular example, however, it holds.

4. So wrote Theodor Lipps in 1880. (Quoted by Chisholm (1966: 79)).

5. I do not think that the conventionalist is entitled to draw this conclusion. In doing so, he is ignoring the possibility that sometimes language users cannot follow the conventions governing the use of a given logical constant without certain ontological assumptions being made.

6. For a recent defense of this position see Ellis (1979).

7. I think that the first to conceive of this possibility were the Neo-Kantians.

8. See for example the first chapter of Nute (1980).

9. There is a ready analogy with the way the Chomsky tradition sees syntactic competence.

10. My own views on the status of logic have been greatly influenced by the teachings of Else Barth. See for example

Barth (1979).

11. This has been stressed more than once by Michael Dummett.

12. This does not mean, however, that I am convinced that for our purposes this really is the best way to proceed. I am aware that if one takes a constructive stand on the ontological status of mathematical objects - as perhaps one should - one just cannot talk the way classical logic suggests. Then, the way of speaking underlying intuitionistic logic is much more suitable.

13. For other ways see II.2.2.

14. C.f. the discussion of possible worlds in chapter II.1.1.

15. A similar result is proved by Gibbard (1981).

16. This could have been done otherwise, but for the present purposes it would not make any difference.

17. See Adams(1975) pp 30 - 33.

18. Harper et. al. (1981) contains most of the important papers on the subject. (This anthology is mentioned in the references under Gibbard (1981).)

19. For a detailed criticism of relevance logic see Copeland (1978).

20. A working knowledge of his work is assumed here.

21. Both examples stem from Adams (1975).

22. See Groenendijk and Stokhof (1982) for further discussion of these points.

23. Perhaps Discourse Representation Theory (see Kamp (1984)) can be seen as a new theory of logical form.

PART II

P O S S I B L E
W O R L D S
S E M A N T I C S

II.1. PRELIMINARIES

II.1.1. Worlds and propositions

Giving a formal analysis of an informal argument is like drawing a cartoon: one has to leave out everything that is unimportant, exaggerate the few things left and when this is properly done the result can be a striking characterization of what is going on. Possible-worlds semantics can best be understood as a certain technique for drawing these cartoons. Ask any two possible-worlds semanticists to analyse the case of the marbles or any other argument involving conditionals, and they will proceed in the same manner.

The first thing they will do is introduce a formal language, some sentences of which represent conditionals. Any of the following meets that requirement.

II.1. DEFINITION. A *language* L has as its vocabulary (i) a number of *atomic sentences*; (ii) three *logical constants*, $\sim$, $\wedge$, and $\rightarrow$; (iii) two *parentheses*, (, and). Given the vocabulary, the set of *sentences* of L is defined as the smallest set X meeting the following conditions: (i) every atomic sentence is an element of X; (ii) if ϕ is an element of X then so is $\sim\phi$; (iii) if ϕ and ψ are elements of X then so are $(\phi \wedge \psi)$ and $(\phi \rightarrow \psi)$.

For the example with the marbles a language with just six atomic sentences is sufficient: 'red in 1', 'red in 2', 'yellow in 1', 'yellow in 2', 'blue in 1' and 'blue in 2'. But the analysis of other arguments may require more atomic sentences, some even infinitely many. That is why in the definition the number of atomic sentences is left undecided.

As for the logical constants, $\sim$ and $\wedge$ are meant as counterparts of negation ('it is not the case that ...') and conjunction ('... and ...') as usual. The languages *L* are, for reasons of economy, not provided with a counterpart of disjunction ('... or ...'). Still, it is convenient to smuggle in the well known wedge 'v' as a metalinguistic abbreviation: '$(\phi \vee \psi)$' will be short for '$\sim(\sim\phi \wedge \sim\psi)$'. More abbreviations will be introduced when needed.

Depending on the kind of example dealt with, '$(\phi \rightarrow \psi)$' must be read as an indicative conditional 'if it is the case that ϕ, then it is the case that ψ', as a subjunctive conditional 'if it were the case that ϕ, then it would be the case that ψ' or as a counterfactual conditional 'if it had been the case that ϕ, then it would have been the case that ψ'. This is not to suggest that everybody thinks that these different locutions all have the same logical properties. (Some possible-worlds semanticists think so, others do not.) It is simply that we will be dealing with only one kind of conditional at a time; so we do not need any more than just one implication sign.

So much for the languages *L*. The second thing that every possible-worlds semanticist will do is introduce a non-empty set of so-called *possible worlds*.

Most possible-worlds semanticists do know what this set could be. Some even think they can explain it independently of the particular argument they are analysing. For them there is just one set of possible worlds, *the* set of possible worlds, and whatever the argument concerned, they always bring this one set of possible worlds into play. According to them the idea of possible worlds is to be understood in a quite literal sense. They are entities of the same sort as the world we happen to live in. They differ from the actual world only in what goes on in them. They may not *actually* exist, since to actually exist is to exist in the actual world, but they nevertheless exist. As David Lewis (1973: 84) puts it: "I believe that there are possible worlds other than the one we happen to inhabit. If an argument is wanted, it is this. It is uncontroversially true that things might be otherwise than they are. I believe, and so do you, that things could have been different in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way they actually are. On the face of it this sentence is an existential quantification. It says that there exist many entities of a certain description, to wit 'ways things could have been'. I believe that things could have been different in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe in the existence of entities that might be called 'ways things could have been'. I prefer to call them 'possible worlds'."

The shift from 'ways things could have been' to 'possible worlds' is by no means as innocent a substitution as Lewis suggests. As Robert Stalnaker (1976: 68) points out: "If possible worlds are ways things might have been, then the actual world ought to be *the ways things are* rather than [as Lewis describes it - F.V.] *I and all my surroundings. The way things are* is a property or a state of the world, not the world itself. The statement that the world is the way it is is true in a sense, but not when read as an identity statement. (Compare: 'the way the world is is the world'.) This is important, since if properties can exist uninstantiated, then *the way the world is* could exist even if a world that is that way did not. One could accept ... that there really are many ways that things could have been - while denying that there exists anything else that is like the actual world."

In other words, there is a difference between a way our world might have been and a world which *is* that way. Lewis's argument shows that there is nothing wrong - or strange - in believing that the former kind of thing exists. But that does not commit ut to the assumption that the latter kind of thing exists. Only the world which is the way things actually are exists, the other ones do not.

So much for the ontological questions surrounding Lewis's possible worlds. On to the more urgent, epistemic questions. After all, even if Lewis is right in maintaining that the worlds he is talking about really exist, it remains that we hardly know anything about them - how many are there, for instance? - and that there is no way of finding out more about them. It is to be feared that our theory of conditionals is going to lack explanatory power if in each and every application we must have recourse to these mysterious entities. And that explanatory power is about the only thing that logical theories cannot miss has been the point of much of part I.

These problems are partly caused by the undue weight of extra baggage being lugged around in order to explain the sort of things we are interested in. The possible worlds involved are full blown working alternatives to the actual world. They can have all its complexity and detail. It seems out of proportion that all of this is really necessary just to explain, say, the example with the three marbles in the two boxes. The only conceivable advantage of introducing this very large set of very complex worlds is that it might be sufficiently large and contain sufficiently complex worlds to cope with all arguments at once. But we do not need one big device that can handle all arguments at once. We can do just as well with many devices, each handling one case. And if this increases the explanatory power of our theory, then it is preferable.

So we will prefer to think of possible worlds in a slightly different manner.[1)] As above, a possible world could be characterized as a-way-things-could-be (or, more precisely, something that is that way) only this time it is a-way-*some*-things-could-be rather than a-way-*all*-things-could-be; it depends on the particular argument we are analysing *which* things matter. This is how Stalnaker (1984: 4) says it: "The term 'possible world' is perhaps misleading for what I have in mind. A set of possible worlds may be a space of relevant alternative possible states of some limited subject matter determined by a context in which some rational activity (deliberation, inquiry, negotiation, conversation) is taking place. Although the kind of abstract account of speech and thought that I will presuppose takes possible worlds for granted, it need not take on the metaphysical burdens which the picturesque terminology suggests. All that is assumed is that agents who think and talk are distinguishing between possibilities, that their so distinguishing is essential to the activities which constitute their thinking and talking, and that we can usefully describe the activities in terms of the possibilities they are distinguishing between."

Sometimes the possible worlds in question may turn out to be very simple sorts of things. Take the case of the marbles: every possible distribution of the marbles over the boxes counts as a possible world. These possible worlds are not very worldlike, all-right, but this is a matter of jargon. How many of these marble-worlds are there? There is room for discussion here. Some might want only six worlds, because that is the number of possible distributions given the information that every marble is in one of the boxes and each box contains at least one marble. (Compare the table in I.1.1.) Others might want many more possible worlds, twenty seven for example, because that is the number of distributions to be reckoned with if this accidental information is left aside. That there is room for discussion here does not make possible-worlds semantics any less viable as one might perhaps be inclined to think. We may disagree about the range of possibilities to be reckoned with. Your eventual analysis may turn out better than mine. Still, as long as we agree that there is such a range of possibilities and that its exact measure is relevant to the case concerned we are both doing possible-worlds semantics, in any case we are well on the way.

Suppose we have chosen a language L and fixed a set W of possible worlds, both suited to the argument we are concerned with. We can now say what exactly we are after. We want to specify which sentences ϕ of L are true at which worlds i in

W. In terms of the following definition: we want to specify which *interpretation* of *L* over *W* is the correct one.

II.2. DEFINITION. Let *L* be a language and *W* a (non-empty) set of possible worlds. An *interpretation of L over W* is a function [ ] that assigns to each sentence ϕ of *L* a subset $[\phi]$ of *W*.
If $i \in [\phi]$ we say that ϕ is *true at* i (*under* the interpretation [ ]).
If $i \notin [\phi]$ we say that ϕ is *false at* i (*under* the interpretation [ ]). □

It is worthwile reading this definition twice - the second time as a further explanation of what possible worlds are. Apparently, no sentence can be both true and false at any possible world i. Moreover, no sentence can be neither true nor false at any possible world i. So, possible worlds are complete in a sense; they may be *small* worlds, depicting a way *some* things could be rather than a way *all* things could be, still each world has to decide all of these things.[2)]

It is common practice among possible-world semanticists to refer to the subsets of the set of possible worlds as *propositions*. In particular, if [ ] is an interpretation of *L* over *W*, then $[\phi]$ is called *the proposition expressed* by ϕ (*under* the interpretation [ ]). In view of the above definition, it will be clear why this is appropriate. You know which set of worlds $[\phi]$ is just in case you know when ϕ is true. Furthermore, the sets $[\phi]$ and $[\psi]$ are identical just in case ϕ and ψ always have the same truth value. What is more natural, then, than calling the set $[\phi]$ the proposition expressed by ϕ, so that we literally mean what we say when we say that two sentences which always have the same truth value express the same proposition. As David Lewis (1973: 46) puts it: "the identification of propositions with sets of worlds captures a good part of the tradition. Propositions so understood are non-linguistic entities ...

One proposition may be expressed by many sentences, in one language or in many, or by non-verbal means of communication; on the other hand there may be propositions that we have no way to express." And he adds an important caveat: "But one part of the tradition about propositions must be given up: propositions understood as sets of worlds cannot serve as the meanings of the sentences that express them, since there are sentences - for instance, all the logical truths - that express the same proposition but do not, in any ordinary sense, have the same meaning." Fortunately, most possible-worlds semanticists are not looking for entities that capture the meaning of a sentence. They are interested in meaning only to the extent that meaning matters to logic. Given their explication of validity only truth conditions really matter. And propositions understood as sets of possible worlds seem just the right entities to capture these.

As said above, we do no want to specify any old interpretation of the language *L* concerned, we want to specify the correct interpretation.

Stalnaker (1968), Lewis (1973), Pollock (1976), Veltman (1976) and Kratzer (1979) - to mention just a few who have developed theories of conditionals within the framework of possible-worlds semantics, all differ as to which interpretation is the correct one. It will be the object of much of the following to compare these theories. In this chapter, however, I am more interested in what they all have in common than in the respects in which they differ. In particular, I am interested in what it is which makes them all possible-worlds theories.

One of such feature is that they all[3] satisfy the following version of the compositionality principle:

The proposition expressed by a compound sentence is uniquely determined by the propositions expressed by its constituent sentences. To put it differently, all theories agree that the correct interpretation is at any rate a *compositional* interpretation in the following sense:

II.3. DEFINITION. Let [ ] be an interpretation of the language L over the set of worlds W. [ ] is *compositional* iff there are functions $F_{\sim}$, $F_{\wedge}$, and $F_{\rightarrow}$ - $F_{\sim}$ a function from propositions in W to propositions in W; $F_{\wedge}$ and $F_{\rightarrow}$ functions from pairs of propositions in W to propositions in W - such that for any ϕ and ψ,

(i) $[\sim\phi] = F_{\sim}([\phi])$;

(ii) $[\phi \wedge \psi] = F_{\wedge}([\phi],[\psi])$;

(iii) $[\phi \rightarrow \psi] = F_{\rightarrow}([\phi],[\psi])$. □

Logically speaking, the version of the compositionality principle given above is equivalent to a criterion of replacement: whenever ϕ and ψ express the same proposition, the replacement of an occurrence of ϕ in a sentence χ by an occurrence of ψ will yield a sentence χ' that has the same truth value as the original χ.[4)]

A logical theory has to supply a well-articulated criterion of replacement - otherwise, it is not worthy of its name. Now, the above criterion might be wrong; it might be that expressing the same proposition is no more a sufficient condition for replacement than having the same truth value is. (Cf. I.2.1.) Suppose, for the sake of argument, that it *is* wrong. Then the immediate question is this: if not their propositional content, what then should two sentences have in common to be substitutable for one other, *salva veritate*? Their meaning? Or should the notion of a proposition be explained differently? Or what? Whatever the answer, it would certainly take us outside the framework of possible-worlds semantics. Therefore, as long as we want to stay within the framework, we are virtually forced to assume that this criterion of replacement holds, and with it the version of the compositionality principle given above.

Assuming that the correct interpretation of a given language L over a given set of worlds W is compositional, our problem boils down to this: what is the correct interpretation of the atomic sentences of L; which function from propositions in W to propositions in W is the right $F_{\sim}$; and which functions from pairs of propositions in W to propositions in W are the right $F_{\wedge}$ and $F_{\rightarrow}$?

As for the atomic sentences we can be short. We will not look for their correct interpretation. We will suppose that their correct interpretation is *given* in advance, determined as it were by our choice of L and W. This is captured formally by the introduction of a function I that tells for each atomic sentence ϕ of L what the proposition $I(\phi)$ expressed by ϕ is. I is called an *atomic interpretation* of L over W.

As for the right $F_{\sim}$ and the right $F_{\wedge}$ we will not spend much time looking for these either. Throughout this part of our study, we will assume, as all[5)] possible-worlds semanticists do, that for any proposition $p \subset W$, $F_{\sim}(p) = W \sim p$, and that for any two propositions $p,q \subset W$ $F_{\wedge}(p,q) = p \cap q$. In other words, a sentence of the form $\sim\phi$ is true at a world i just in case ϕ is false at i. And a sentence of the form $(\phi \wedge \psi)$ is true at i just in case both ϕ and ψ are true at i.

Finding the right $F_{\rightarrow}$ will take considerably more time. Until further notice every function F from pairs of propositions in W to propositions in W - henceforth: every *connector* with W as its *domain* - counts as a possible candidate.

Putting the above considerations together, we state

II.4. DEFINITION. Let L be a language, W a set of possible worlds, I an atomic interpretation of L over W, and F a connector with W as its domain. Consider any interpretation [ ] of L over W.

[ ] *conforms to* I iff for every atomic sentence ϕ $[\phi] = I(\phi)$.

[ ] is *standard* iff for any two sentences ϕ and ψ

(i) $[\sim\phi] = W \sim [\phi]$;

(ii) $[\phi \wedge \psi] = [\phi] \cap [\psi]$

[ ] is *based on* F iff for any two sentences ϕ and ψ
$[\phi \rightarrow \psi] = F([\phi],[\psi])$. □

We still do not know what the correct interpretation of a given language L over a given set of possible worlds W is. All we know is that it will be a standard interpretation conforming to a pregiven atomic interpretation, and that it will be based on an as yet unidentified connector. And that is about the least a possible-worlds semanticist can and must say about it.

II.1.2. Logical notions

II.5. DEFINITION.
(i) Let F be a connector with domain W. A set Δ of sentences (of a given language L[6]) is *satisfiable by* F iff there is a standard interpretation [ ] based on F such that for some $i \in W$, $i \in [\phi]$ for every $\phi \in \Delta$. An argument Δ/ϕ (i.e. the argument with premises Δ and conclusion ϕ) is *valid on* F iff $\Delta \cup \{\sim\phi\}$ is not satisfiable by F. A sentence ϕ is *valid on* F iff the argument $\emptyset/\phi$ is valid on F.
(ii) Let K be a class of connectors. An argument Δ/ϕ is *valid within* K iff Δ/ϕ is valid on each $F \in K$. A sentence ϕ is *valid within* K iff $\emptyset/\phi$ is valid within K. □

Of course, we would have proceeded differently, if we had already identified the right connectors. If we had known, for every domain W of possible worlds, which connector is best to be taken as the basis for the interpretation of our languages over W, then we would have taken the class of these connectors as the starting point for an absolute definition of logical validity, whereas now this notion remains relative to arbitrary classes of connectors.

II.6. DEFINITION. Let $\mathcal{L}$ be a set of arguments. $\mathcal{L}$ is a *logic* iff the following holds.

- Identity : $\phi/\phi \in \mathcal{L}$ for every sentence ϕ.
- Augmentation : if $\Delta/\phi \in \mathcal{L}$, then $\Gamma/\phi \in \mathcal{L}$ for every $\Gamma \supset \Delta$.
- Cut : if $\Delta,\Gamma/\phi \in \mathcal{L}$ and $\Delta/\psi \in \mathcal{L}$ for every $\psi \in \Gamma$, then $\Delta/\phi \in \mathcal{L}$.
- Substitution : if $\Delta/\phi \in \mathcal{L}$ then $\Delta'/\phi' \in \mathcal{L}$, where Δ'/ϕ' is the result of substituting an arbitrary sentence ψ' for some atomic sentence ψ at all places where ψ occurs in ϕ and in the sentences of Δ.
- Replacement : if $\phi/\psi \in \mathcal{L}$, $\psi/\phi \in \mathcal{L}$, and $\Delta/\chi \in \mathcal{L}$, then $\Delta'/\chi' \in \mathcal{L}$ where Δ'/χ' is the result of substituting ϕ for ψ at one or more places where ψ occurs in χ or in the sentences of Δ.
- Elimination $\wedge$: if $\Delta/\phi \wedge \psi \in \mathcal{L}$, then $\Delta/\phi \in \mathcal{L}$ and $\Delta/\psi \in \mathcal{L}$.
- Introduction $\wedge$: if $\Delta/\phi \in \mathcal{L}$ and $\Delta/\psi \in \mathcal{L}$, then $\Delta/\phi \wedge \psi \in \mathcal{L}$.
- Elimination $\sim$: if $\Delta/{\sim}\phi \in \mathcal{L}$ and $\Delta/\phi \in \mathcal{L}$, then $\Delta/\psi \in \mathcal{L}$.
- Introduction $\sim$: if $\Delta,\phi/{\sim}\phi \in \mathcal{L}$ then $\Delta/{\sim}\phi \in \mathcal{L}$.
- Elimination $\sim\sim$: if $\Delta/{\sim}{\sim}\phi \in \mathcal{L}$ then $\Delta/\phi \in \mathcal{L}$. □

II.7. DEFINITION. Let K be a class of connectors. $\mathcal{L}_K$, the *logic determined by* K, is the set of all arguments that are valid within K. □

Maybe we have been a little hasty in calling $\mathcal{L}_K$ a logic.

II.8. PROPOSITION. Let K be a class of connectors. $\mathcal{L}_K$ is a logic in the sense of definition 6.

PROOF. Routine. □

What about the converse of this proposition? Is every logic determined by some class of connectors? Rather surprisingly, the answer is no. Some logics are in this sense *incomplete*.

II.9. DEFINITION. Let $\mathcal{L}$ be a logic.

(i) $\mathcal{L}$ is *complete* iff there is a class K of connectors such that $\mathcal{L} = \mathcal{L}_K$.[7)]

(ii) The *range* of $\mathcal{L}$ is the class $K_{\mathcal{L}}$ of all connectors on which every argument belonging to $\mathcal{L}$ is valid. □

II.10. PROPOSITION. Let $\mathcal{L}$ be a logic.

(i) $\mathcal{L} \subset \mathcal{L}_{(K_{\mathcal{L}})}$.

(ii) $\mathcal{L}$ is complete iff $\mathcal{L}_{(K_{\mathcal{L}})} = \mathcal{L}$.

PROOF. Obvious. □

So, if the logic $\mathcal{L}$ is complete, the logic determined by the range of $\mathcal{L}$ is $\mathcal{L}$ itself. If $\mathcal{L}$ is incomplete, the logic determined by the range of $\mathcal{L}$ is stronger than $\mathcal{L}$ suggests.

II.11. DEFINITION. Let K be a class of connectors. Call a set $\mathbf{A}$ of arguments *characteristic of* K iff a connector F belongs to K just in case every argument in $\mathbf{A}$ is valid on F. K is *characterizable* iff there is a set of arguments which is characteristic of K. □

II.12. PROPOSITION. Let K be a class of connectors.

(i) $K \subset K_{(\mathcal{L}_K)}$.

(ii) K is characterizable iff $K = K_{(\mathcal{L}_K)}$.

PROOF. Omitted. □

(i) says that every class of connectors forms part of the range of its logic; (ii) adds that for characterizable classes of connectors the converse is also true.

Obviously no class K is characterizable unless it is closed under isomorphism: if the connector F belongs to K, and the connector G is an isomorphic copy of F, then G must also belong to K.[8)] As we will see, however, even among the classes of connectors that fulfill this condition, there are many non-characterizable ones. Our formal languages just lack the power needed to express all structural properties that connectors can have.

In the following we will be much more concerned with questions of the form 'what is the logic determined by this particular class of connectors?' than with questions of the form 'by which class of connectors, if any, is this particular logic determined?'. Consequently, there will be a much higher chance of running into a non-characterizable class of connectors than of running into an incomplete logic. Actually, even if we were concerned with the latter kind of question, there would be little chance of meeting up with one of these. They are generally well hidden away; you really do have to look for them. The only examples I know of are in Nute (1978), who adapted some incomplete modal logics to conditionals.[9)]

II.13. DEFINITION. Let $\mathcal{L}$ be any logic. $\mathcal{L}$ is *compact* iff for every $\Delta/\phi \in \mathcal{L}$ there is some finite $\Gamma \subset \Delta$ such that $\Gamma/\phi \in \mathcal{L}$. □

II.14. PROPOSITION. Let $\mathfrak{M}$ be the weakest logic (i.e. the smallest set of arguments fulfilling the conditions laid down in definition 6). $\mathfrak{M}$ is compact.

PROOF. Induction on $\mathfrak{M}$. □

II.15. DEFINITION. Let $\mathcal{L}$ be any logic. A set Δ of sentences is $\mathcal{L}$-*consistent* iff there is no sentence ϕ such that both $\Delta/\phi \in \mathcal{L}$ and $\Delta/{\sim}\phi \in \mathcal{L}$. Δ is *maximally* $\mathcal{L}$-consistent iff Δ is $\mathcal{L}$-consistent while no proper extension Γ of Δ is $\mathcal{L}$-consistent. $\mathcal{L}$ is *consistent* iff $\emptyset$ is $\mathcal{L}$-consistent. □

In view of proposition 8, $\mathfrak{M}$ is consistent.

II.16. DEFINITION.

(i) Let $\mathcal{L}$ be any logic. An interpretation [ ] over a given set of worlds W *verifies* $\mathcal{L}$ iff for every $i \in W$, $\{\phi \mid i \in [\phi]\}$ is $\mathcal{L}$-consistent.

(ii) $\mathcal{L}$ is *canonical* iff every standard, compositional interpretation that verifies $\mathcal{L}$ can be based on some connector $F \in K_{\mathcal{L}}$. □

$K_{\mathfrak{M}}$ is the class of *all* connectors. By definition, every compositional interpretation is based on *some* connector. Therefore, $\mathfrak{M}$ is canonical.

$\mathfrak{M}$ is compact, consistent and canonical. We will now prove that every logic with these properties is complete.

II.17. LEMMA (*Lindenbaum's Lemma*). If $\mathcal{L}$ is compact, every $\mathcal{L}$-consistent Δ has a maximal $\mathcal{L}$-consistent extension.

PROOF. Usually, Lindenbaum's lemma is stated for systems of deduction - not for the kind of logics introduced in definition 6. This hardly matters to the proof, though: here the compactness of $\mathcal{L}$ plays much the same role as the finite length of deductions does in the usual case. □

II.18. LEMMA. Let Δ be maximally $\mathcal{L}$-consistent. Then
(i) $\phi \in \Delta$ iff $\Delta/\phi \in \mathcal{L}$;
(ii) $\sim\phi \in \Delta$ iff $\phi \notin \Delta$;
(iii) $\phi \wedge \psi \in \Delta$ iff $\phi \in \Delta$ and $\psi \in \Delta$.

PROOF. Standard. □

II.19. THEOREM. Every consistent, compact and canonical logic is complete.

PROOF. Let $\mathcal{L}$ be consistent, compact and canonical. Let W be given by $W = \{\Delta \mid \Delta$ is maximally $\mathcal{L}$-consistent$\}$. Since $\mathcal{L}$ is consistent and compact $W \neq \emptyset$. Consider the interpretation [ ] over W defined by $[\phi] = \{\Delta \in W \mid \phi \in \Delta\}$.

Using lemma 18, it can straightforwardly be proved that [ ] is a standard interpretation.

Next we show that [ ] is compositional. It is sufficient to show that if $[\phi] = [\phi']$ and $[\psi] = [\psi']$, then $[\phi \rightarrow \psi] = [\phi' \rightarrow \psi']$. If $[\phi] = [\phi']$, then $\phi/\phi' \in \mathcal{L}$ and $\phi'/\phi \in \mathcal{L}$. And if $[\psi] = [\psi']$, then $\psi/\psi' \in \mathcal{L}$ and $\psi'/\psi \in \mathcal{L}$. Applying the replacement rule of definition 6, we find that $\phi \rightarrow \psi/\phi' \rightarrow \psi' \in \mathcal{L}$ and $\phi' \rightarrow \psi'/\phi \rightarrow \psi \in \mathcal{L}$. This means that $[\phi \rightarrow \psi] = [\phi' \rightarrow \psi']$.

So, [ ] is a standard, compositional interpretation that verifies $\mathcal{L}$. Since $\mathcal{L}$ is canonical this means that there is a connector $F \in K_{\mathcal{L}}$ on which [ ] can be based.

We are now ready to show that $\mathcal{L}$ is complete. In view of proposition 10 it is sufficient to show that $\mathcal{L}_{(K_{\mathcal{L}})} \subset \mathcal{L}$. Suppose $\Delta/\phi \notin \mathcal{L}$. Then $\Delta' = \Delta \cup \{\sim\phi\}$ is $\mathcal{L}$-consistent. Since $\mathcal{L}$ is compact, Δ' has a maximal $\mathcal{L}$-consistent extension Γ. Given the definition of W, $\Gamma \in W$. Given the definition of [ ], $\Gamma \in [\phi]$ for every $\phi \in \Gamma$. This shows that Γ is satisfiable by F. Hence, since $F \in K_{\mathcal{L}}$, Δ/ϕ is not valid on some connector in $K_{\mathcal{L}}$. In other words, $\Delta/\phi \notin \mathcal{L}_{(K_{\mathcal{L}})}$. □

As a first application, we see that $\mathfrak{M}$ is complete.

It will appear that theorem 19 provides a very fruitful method for proving completeness results. In fact, this method works so well, that one may wonder whether it is generally applicable: is every consistent, compact and complete logic canonical?

The question has not yet been settled. No example is known showing that the answer is no, but neither is there a proof showing that it is yes. (The same holds for modal logic, although some special results in Fine (1974) and Van Benthem (1980) suggest that for compact and consistent modal logics, completeness and canonicity do indeed coincide.)

By now, you will be accustomed to the rather unorthodox way in which the word 'logic' is used in this chapter. I even hope it has become clear that the way it is used is not that unorthodox after all. In fact, this way we cover two more or less standard usages.

If we had followed common practice we would have had to make a choice, using the word 'logic' either in connection with classes of connectors or in connection with systems of proof. In the former case, custom would have dictated this definition:

- Let K be a class of connectors. $\mathcal{L}^{K} = \{\phi \mid \phi$ is valid within $K\}$.

So we would have thought of the logic determined by a class K of connectors as the set of sentences valid within K rather than as the set of arguments valid within K. Note, however, that once you know what $\mathfrak{L}_K$ is you also know what $\mathfrak{L}^K$ is. Moreover, if $\mathfrak{L}_K$ is compact, the converse holds as well. To put it differently, as long as $\mathfrak{L}_K$ is compact, the difference between $\mathfrak{L}_K$ and $\mathfrak{L}^K$ is merely one of presentation.

In the latter case, we would have specified a number of proof systems. The definition of a logic would then have been something like this.

- Let Σ be a proof system. $\mathfrak{L}^\Sigma = \{\Delta/\phi \mid \phi$ is provable from Δ in $\Sigma\}$.

The connection between our logics and the logics generated by proof systems may have become clear from the discussion of the logic $\mathfrak{M}$. $\mathfrak{M}$ is resursively enumerable, as is obvious from its definition, but the definition does not supply a proof system enumerating it. What it supplies is an inventory of the properties that any such system must have. Fortunately, this inventory is so detailed that a suitable proof system can easily be extracted. (A Gentzen-like calculus of sequents suggests itself.)

It will be clear that these remarks apply to many logics other than $\mathfrak{M}$.

II.2. DELINEATIONS

II.2.1. Constraints on consequents

Let $\mathfrak{C}$ be the weakest logic $\mathfrak{L}$ with the following properties:
Free Deduction: if $\Delta,\phi/\psi \in \mathfrak{L}$ then $\Delta/\phi \rightarrow \psi \in \mathfrak{L}$;
Modus Ponens : if $\Delta/\phi \rightarrow \psi \in \mathfrak{L}$ and $\Delta/\phi \in \mathfrak{L}$ then $\Delta/\psi \in \mathfrak{L}$.

$\mathfrak{C}$ is better known as *classical* logic, and it is not only the weakest logic with these properties but also the strongest one: every consistent logic that includes $\mathfrak{C}$ is equal to $\mathfrak{C}$.

Actually, most logicians think $\mathfrak{C}$ is much too strong. Or rather, many think it is much too strong for indicative conditionals, and all agree that it is much too strong for counterfactuals.

Let us concentrate on the latter case. Perhaps the quickest way to see that $\mathfrak{C}$ is much too strong for counterfactuals is this.

II.20. PROPOSITION. Let $\mathfrak{L}$ be a logic. $\mathfrak{L}$ is closed under Free Deduction iff for all ϕ,ψ both $\sim\phi/\phi \rightarrow \psi \in \mathfrak{L}$ and $\psi/\phi \rightarrow \psi \in \mathfrak{L}$.

PROOF. Suppose $\mathfrak{L}$ is closed under Free Deduction. By Identity and Augmentation we have that $\psi,\phi/\psi \in \mathfrak{L}$. Hence, $\psi/\phi \rightarrow \psi \in \mathfrak{L}$. By Identity and Augmentation, both $\sim\phi,\phi/\phi \in \mathfrak{L}$ and $\sim\phi,\phi/\sim\phi \in \mathfrak{L}$. By Elimination $\sim$, it follows that $\sim\phi,\phi/\psi \in \mathfrak{L}$. So, by Free Deduction $\sim\phi/\phi \rightarrow \psi \in \mathfrak{L}$.

Remark: a neat shorthand notation suggests itself; in this notation the preceding argument becomes:

$$\dfrac{\dfrac{\dfrac{\phi/\phi}{\sim\phi,\phi/\phi}\text{Aug}\qquad\dfrac{\sim\phi/\sim\phi}{\sim\phi,\phi/\sim\phi}\text{Aug}}{\sim\phi,\phi/\psi}\text{E}}{\sim\phi/\phi\rightarrow\psi}\text{FD}$$

For the converse, we first note that every logic $\mathcal{L}$ is closed under the following rule:

(*) If $\Delta,\phi/\psi \in \mathcal{L}$ and $\Delta,\sim\phi/\psi \in \mathcal{L}$ then $\Delta/\psi \in \mathcal{L}$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Delta,\phi/\psi}{\Delta,\sim\psi,\phi/\psi}\text{Aug}\quad\dfrac{\sim\psi/\sim\psi}{\Delta,\sim\psi,\phi/\sim\psi}\text{Aug}}{\Delta,\sim\psi,\phi/\sim\phi}\text{E}}{\Delta,\sim\psi/\sim\phi}\text{I}\sim\qquad\dfrac{\dfrac{\dfrac{\Delta,\sim\phi/\psi}{\Delta,\sim\psi,\sim\phi/\psi}\text{Aug}\quad\dfrac{\sim\psi/\sim\psi}{\Delta,\sim\psi,\sim\phi/\sim\psi}\text{Aug}}{\Delta,\sim\psi,\sim\phi/\sim\sim\phi}\text{E}\sim}{\Delta,\sim\psi/\sim\sim\phi}\text{I}\sim}{\Delta,\sim\psi/\sim\sim\psi}\text{E}\sim}{\Delta/\sim\sim\psi}\text{I}\sim}{\Delta/\psi}\text{E}\sim\sim$$

Using (*), we can complete the proof as follows:

$$\dfrac{\dfrac{\Delta,\phi/\psi\qquad\dfrac{\psi/\phi\rightarrow\psi}{\Delta,\phi,\psi/\phi\rightarrow\psi}\text{Aug}}{\Delta,\phi/\phi\rightarrow\psi}\text{Cut}\qquad\dfrac{\sim\phi/\phi\rightarrow\psi}{\Delta,\sim\phi/\phi\rightarrow\psi}\text{Aug}}{\Delta/\phi\rightarrow\psi}(*)$$

□

A counterfactual $\phi \rightarrow \psi$ is normally uttered in a context where the antecedent ϕ is known to be false, and our theory will have to explain why this is so. However, if this theory dubs $\sim\phi/\phi \rightarrow \psi$ as a valid argument form, it will at best explain why this is *not* so. Indeed, what could possibly be the point of uttering a sentence beginning with 'if it had been the case that ϕ' in such a context if its end 'then it would have been the case that ψ' could be chosen at random.

Similar remarks apply to $\psi/\phi \rightarrow \psi$. After all, it often happens that a sentence ψ is true - regrettably so - and that we want to point out to those responsible for this that 'it would not have been the case that ψ if only ...'. But what could be the use of such a statement if everybody was always justified in replying that 'ψ would have been the case anyway'?

We cannot accept Free Deduction, but we may want to keep some restricted form of this principle. Free Deduction says that if you want to show that $\phi \rightarrow \psi$ follows from Δ, it is sufficient to show that ψ follows from Δ and ϕ. Now clearly, if $\phi \rightarrow \psi$ is a counterfactual, you cannot just *add* ϕ to the premises Δ, supposing as it were that in addition to the sentences in Δ ϕ *is* also true. In such a case you will have to start with a counterfactual hypothesis 'suppose that ϕ *had been* true', and it may very well be that some of the sentences in Δ are incompatible with this hypothesis. What if $\sim\phi \in \Delta$, for instance? Would $\sim\phi$ still have been true if ϕ had been true? On the other hand, there is nothing wrong in appealing to *given* counterfactual consequents of ϕ in such a case. Indeed, the least one can say is that in order to show that the counterfactual $\phi \rightarrow \psi$ follows from Δ, it is sufficient to show that ψ follows from ϕ and $\{\chi \mid \phi \rightarrow \chi \in \Delta\}$.

In the next definition Δ^{ϕ} is used as an abbreviation of $\{\chi \mid \phi \rightarrow \chi \in \Delta\}$.

II.21. DEFINITION. A *conditional* logic is any logic $\mathcal{L}$ with the following property:
Restricted Deduction: if $\Delta^{\phi},\phi/\psi \in \mathcal{L}$ then $\Delta/\phi \rightarrow \psi \in \mathcal{L}$. □

I would not call an arrow $\rightarrow$ that does not even allow this restricted form of hypothetical reasoning a conditional operator.

II.22. DEFINITION. Let F be a connector with domain W. F is *conclusive* iff for every $i \in W$ and $p,q,r \subset W$

(i) $i \in F(p,p)$;

(ii) if $i \in F(p,q)$ and $q \subset r$, then $i \in F(p,r)$;

(iii) if $i \in F(p,q)$ and $i \in F(p,r)$, then $i \in F(p,q \cap r)$.

F is *strongly conclusive* if (iii) can be strengthened to

(iii)' if $Q \neq \emptyset$ and $i \in F(p,q)$ for every $q \in Q$, then $i \in F(p,\cap Q)$. □

II.23. PROPOSITION. Let $\mathbf{B}$ be the weakest conditional logic. Let F be any connector. The following three statements are equivalent

(i) $F \in K_{\mathbf{B}}$;

(ii) F validates all arguments of the form

Conditional Identity (CI) : $\emptyset/\phi \to \phi$,

Weakening the Consequent (CW) : $\phi \to \psi/\phi \to (\psi \vee \chi)$,

Conjunction of Consequents (CC): $\phi \to \psi, \phi \to \chi/\phi \to (\psi \wedge \chi)$;

(iii) F is conclusive.

PROOF.

(i) implies (ii). This is proved by showing that the arguments mentioned under (a), (b) and (c) belong to $\mathbf{B}$.

(ii) implies (iii). This is obvious.

(iii) implies (i). Note first that $\mathbf{B}$ is compact. We must show that if $\Delta/\phi \in \mathbf{B}$, Δ/ϕ is valid on F. This is proved by induction on $\mathbf{B}$. In view of proposition 8, the only interesting case is that of Restricted Deduction: Suppose $\Delta/\phi \to \psi \in \mathbf{B}$ in virtue of the fact that $\Delta^{\phi},\phi/\psi \in \mathbf{B}$. By compactness there is some finite $\Gamma \subset \Delta$ such that $\Gamma^{\phi},\phi/\psi \in \mathbf{B}$; given the induction hypothesis, we may assume that $\Gamma^{\phi},\phi/\psi$ is valid on F, which means that for every standard interpretation [ ] based on F,

$$\cap\{[\chi] \mid \chi \in \Gamma^{\phi} \cup \{\phi\}\} \subset [\psi]$$

Now, consider any standard interpretation [ ] based on F, and any i in the domain of F such that $i \in [\theta]$ for every $\theta \in \Delta$. Then $i \in [\phi \to \chi]$ for every $\chi \in \Gamma^{\phi}$. In other words, $i \in F([\phi],[\chi])$ for every $\chi \in \Gamma^{\phi}$. Since F is conclusive, $i \in F([\phi],[\phi])$. Hence,

$i \in F([\phi],[\chi])$ for every $\chi \in \Gamma^{\phi} \cup \{\phi\}$.

Since F is conclusive, and $\Gamma^{\phi} \cup \{\phi\}$ is finite, this means that

$i \in F([\phi], \cap\{[\chi] \mid \chi \in \Gamma^{\phi} \cup \{\phi\}\})$.

Since F is conclusive, and $\cap\{[\chi] \mid \chi \in \Gamma^{\phi} \cup \{\phi\}\} \subset [\psi]$, we get

$i \in F([\phi],[\psi])$.

That is, $i \in [\phi \rightarrow \psi]$.

This shows that $\Delta/\phi \rightarrow \psi$ is valid on F. □

II.24. THEOREM. $\mathcal{B}$ is complete.

PROOF. By theorem 19 it is sufficient to show that $\mathcal{B}$ is canonical. Let [ ] be any standard, compositional interpretation verifying $\mathcal{B}$. Assume that $\mathcal{W}$ is the set of worlds over which [ ] is defined. Consider the connector F with domain $\mathcal{W}$ defined as follows: $i \in F(p,q)$ iff
either (i) there is no ϕ such that $[\phi] = p$;
or (ii) there are ϕ,ψ such that $[\phi] = p$, $[\psi] \subset q$, and $i \in [\phi \rightarrow \psi]$.

First, we prove that [ ] can be based on F. Obviously, if $i \in [\phi \rightarrow \psi]$ then $i \in F([\phi],[\psi])$. Now assume that $i \in F([\phi],[\psi])$. Then there are ϕ' and ψ' such that $[\phi'] = [\phi],[\psi'] \subset [\psi]$, and $i \in [\phi' \rightarrow \psi']$. Since [ ] is standard $[\psi'] = [\psi'] \cap [\psi] = [\psi' \wedge \psi]$. Since [ ] is compositional, it follows that $i \in [\phi \rightarrow (\psi' \wedge \psi)]$. Since [ ] verifies $\mathcal{B}$, and $\phi \rightarrow (\psi' \wedge \psi)/\phi \rightarrow \psi \in \mathcal{B}$, this implies that $i \in [\phi \rightarrow \psi]$.

Next we show that $F \in K_{\mathcal{B}}$. By the preceding proposition this means that we must show that F is conclusive.

(i) For every i and p, $i \in F(p,p)$. The only interesting case is the case in which $p = [\phi]$ for some ϕ. We have seen above that $\emptyset/\phi \rightarrow \phi \in \mathcal{B}$. Since [ ] verifies $\mathcal{B}$, it follows that $i \in [\phi \rightarrow \phi]$. Hence $i \in F([\phi],[\phi])$.

(ii) If $i \in F(p,q)$ and $q \subset r$, then $i \in F(p,r)$. This is an immediate consequence of the definition of F.

(iii) If $i \in F(p,q)$ and $i \in F(p,r)$, then $i \in F(p,q \cap r)$. Again, the only interesting case is the case in which

$p = [\phi]$ for some ϕ. Suppose $i \in F([\phi],q)$. Then there is a sentence ψ such that $[\psi] \subset q$ and $i \in [\phi \rightarrow \psi]$. Likewise, if $i \in F([\phi],r)$ there is a sentence χ such that $[\chi] \subset r$ and $i \in [\phi \rightarrow \chi]$. Clearly, $[\psi \wedge \chi] = [\psi] \cap [\chi] \subset q \cap r$. Furthermore, since [ ] verifies $\mathbf{B}$ and $\phi \rightarrow \psi, \phi \rightarrow \chi / \phi \rightarrow (\psi \wedge \chi) \in \mathbf{B}$, we have that $i \in [\phi \rightarrow (\psi \wedge \chi)]$. So, $i \in F([\phi], q \cap r)$.

(i), (ii), and (iii) prove that F is conclusive. From this and the fact that [ ] can be based on F it is clear that $\mathbf{B}$ is canonical. □

Notice that the above proof shows that $\mathbf{B}$ can also be described as the weakest logic containing all arguments of the form CI, CC and CW.

II.25. DEFINITION. Let F be a connector with domain W. Consider the function C that assigns to each $i \in W$ and each $p \subset W$ the set $C_i(p)$ of propositions given by

$q \in C_i(p)$ iff $i \in F(p,q)$.

C is called the *consequence function derived from* F. If $q \in C_i(p)$, we say that q is a consequent of the antecedent p at i.

Likewise, the *antecedence function derived from* F is the function A that assigns to each $i \in W$ and $p \subset W$ the set $A_i(p)$ of propositions given by

$q \in A_i(p)$ iff $i \in F(q,p)$.

If $q \in A_i(p)$, q is called an antecedent of the consequent p at i. □

The above definitions work both ways. The consequence function derived from F uniquely determines F, and so does the antecedence function derived from F. Therefore, if we want to, we can choose to begin with the consequence or antecedence functions, deriving the connectors from them instead of the other way around.[10)]

The notion of a consequence function will prove to be of great heuristic value. In particular it will help us get to grips with conclusive connectors.

Let F be any connector. Let C be the consequence function derived from F. Saying that F is conclusive amounts to saying that for every world i and proposition p, the set $C_i(p)$ - i.e. the set of consequents of p at i - constitutes a *filter* (that contains p as an element): every r that includes the intersection $q_1 \cap q_2 \cdots \cap q_n$ of finitely many consequents $q_1, \ldots, q_n$ of p at i is again a consequent of p at i. Saying that F is strongly conclusive amounts to saying that each $C_i(p)$ constitutes a *principal filter* (that contains p as an element). In this case every r which includes the intersection of arbitrarily many consequents of p at i is again a consequent of p at i. In particular, $\cap C_i(p)$, i.e. the intersection of all consequents of p at i, is a consequent of p at i. $\cap C_i(p)$ is, in fact, the strongest consequent of p at i: if $q \in C_i(p)$ then $\cap C_i(p) \subset q$.

Thus, strongly conclusive connectors give rise to the following.

II.26. DEFINITION. A *selection-function*[11] for a given set W of possible worlds is any function S that assigns to each world i in W and each proposition p in W a subset $S_i(p)$ of p. If $j \in S_i(p)$, we say that j is *relevant* to p at i. □

II.27. DEFINITION. Let [ ] be a standard interpretation over a given set of worlds W, and let S be a selection function for W. [ ] is *based on* S iff for any two sentences ϕ and ψ,

$i \in [\phi \to \psi]$ iff $S_i([\phi]) \subset [\psi]$. □

What these definitions tell is that a conditional $\phi \to \psi$ is true at a given world i just in case the consequent ψ is true at all $[\phi]$-worlds that are in some sense relevant to $[\phi]$ at i. Of course, since we do not know in which sense these $[\phi]$-worlds are relevant to $[\phi]$ at i, this can hardly count as a theory. But at least it is a start. And it is precisely what is conveyed by the notion of a strongly conclusive connector.

II.28. DEFINITION. Let F be a connector with domain W, and let S be a selection function for W. F is *interchangeable* with S iff for every $i \in W$ and $p,q \subset W$,

$i \in F(p,q)$ iff $S_i(p) \subset q$ □

If F and S are interchangeable, they give the same meaning to $\rightarrow$.

II.29. THEOREM. A connector F is interchangeable with a selection function iff F is strongly conclusive.

PROOF. Suppose F is interchangeable with S. Let i be any element of W and p be any proposition in W. Then we have

(i) $i \in F(p,p)$, because $S_i(p) \subset p$.

(ii) if $i \in F(p,q)$, then $S_i(p) \subset q$. If, in addition, $q \subset r$, then $S_i(p) \subset r$. If $S_i(p) \subset r$, then $i \in F(p,r)$. So, if $i \in F(p,q)$ and $q \subset r$, then $i \in F(p,r)$.

(iii) if $Q \neq \emptyset$ and $i \in F(p,q)$ for every $q \in Q$, then $S_i(p) \subset q$ for every $q \in Q$. This means that $S_i(p) \subset \cap Q$. Hence, $i \in F(p,\cap Q)$.

This shows that F is strongly conclusive.

To prove the converse, suppose that F is strongly conclusive. Let S for every i and p in the domain W of F be defined by:

$S_i(p) = \cap\{r \mid i \in F(p,r)\}$.

Note that $i \in F(p,S_i(p))$. Hence, $S_i(p) \subset q$ iff $i \in F(p,q)$, which means that F is interchangeable with S. □

As we saw earlier (proposition 23), the class of conclusive connectors is characterized by three argument forms: conditional identity, CI; weakening the consequent, CW; and conjunction of consequents, CC.

The obvious question now arises as to what arguments we must add to this list in order to get a characterization of the class of strongly conclusive connectors. Unfortunately, it cannot be answered. The class of strongly conclusive connectors is not characterizable. Since our languages admit only finite conjunctions, it is impossible to express that the set of consequents of a given antecedent is closed under arbitrary (possibly infinite) intersections. We will need one more definition, and a lemma before we can demonstrate this.

II.30. DEFINITION. Let F be a connector with domain W. Consider the set U of ultrafilters on W. Let * be the function from propositions in W to propositions in U given by

$p^* = \{u \in U \mid p \in u\}$.

Let G be a connector with domain U. G is an *ultrafilter extension* of F iff for every $p,q \subset W$,

$G(p^*,q^*) = (F(p,q))^*$. □

II.31. LEMMA. Let G be an ultrafilter extension of F. Every set of sentences satisfiable by F is satisfiable by G.

PROOF. Let W, U, F, * and G be as above. Let I be any atomic interpretation over W, and consider the atomic interpretation J over U that is related to I as follows:

$J(\phi) = I(\phi)^*$

We show, by induction on the complexity of ϕ, that the interpretation [ ] conforming to I and based on F and the interpretation $[\![\ ]\!]$ conforming to J and based on G are likewise related, i.e.

for every sentence ϕ, $[\![\phi]\!] = [\phi]^*$.

As for the case in which $\phi = \sim\psi$, observe that $(W \sim p)^* = U \sim p^*$. The case in which $\phi = \psi \wedge \chi$ follows smoothly once it has been proved that $(p \cap q)^* = p^* \cap q^*$.

Here is the proof for the case in which $\phi = (\psi \rightarrow \chi)$. By definition, $[\![\psi \rightarrow \chi]\!] = G([\![\psi]\!],[\![\chi]\!])$. By the induction hypothesis, $G([\![\psi]\!],[\![\chi]\!]) = G([\psi]^*,[\chi]^*)$. By definition, $G([\psi]^*,[\chi]^*) = (F([\psi],[\chi]))^* = [\phi \rightarrow \psi]^*$.

The lemma now follows immediately from the observation that if $i \in [\phi]$ for every $\phi \in \Delta$, then the ultrafilter u generated by $\{i\}$ has the property that $u \in [\![\phi]\!]$ for every $\phi \in \Delta$. □

II.32. COROLLARY. A class of connectors is characterizable only if its complement is closed under ultrafilter extensions.

II.33. THEOREM. Let F be any conclusive connector. There is an ultrafilter extension G of F such that G is strongly conclusive.

PROOF. Suppose that W is the domain of F, and let U and * be defined as above. We will use 'P', 'Q' and 'R' as variables ranging over the propositions in U, and 'p', 'q', 'r' as variables ranging over the propositions in W.

Consider for every $u \in U$ and $p \subset W$ the set $\sigma_u(p) = \{r \mid F(p,r) \in u\}$.

Note:

(i) Since F is conclusive, $F(p,p) = W$ for any p. Therefore $F(p,p) \in u$ for any $u \in U$, which means that $p \in \sigma_u(p)$ for any p and $u \in U$.

(ii) Since F is conclusive, $F(p,q) \cap F(p,r) \subset F(p,q \cap r)$. Since any $u \in U$ is a filter, this means that if $F(p,q) \in u$ and $F(p,r) \in u$ then $F(p,q \cap r) \in u$. In other words, if $q \in \sigma_u(p)$ and $r \in \sigma_u(p)$ then $q \cap r \in \sigma_u(p)$.

(iii) Likewise, we see that if $q \in \sigma_u(p)$ and $q \subset r$, then $r \in \sigma_u(p)$.

Now consider the selection function S for U defined by

$$S_u(P) = \begin{cases} \{v \in U \mid \sigma_u(p) \subset v\} & \text{if } P = p^* \\ \emptyset, & \text{otherwise} \end{cases}$$

Since * is injective, this is unambiguous. By (i) above we have that if $P = p^*$, $S_u(P) \subset P$. So, S is indeed a selection function.

Finally, let G be the connector determined by S, i.e. set

$u \in G(P,Q)$ iff $S_u(P) \subset Q$.

By proposition 29, G is strongly conclusive.

So it remains to be proved that G is an ultrafilter extension of F. Suppose $u \notin (F(p,q))^*$. By definition this means that $F(p,q) \notin u$. In other words, $q \notin \sigma_u(p)$. Now as we have seen above ((ii), (iii)), $\sigma_u(p)$ is a filter. Therefore, the latter is equivalent to:

There is some ultrafilter $v \in U$ such that $\sigma_u(p) \subset v$ and $W \sim q \in v$.

This in turn means that there is some $v \in S_u(p^*)$ such that

$v \notin q^*$. Which, by the interchangeability of S and G, amounts to $u \notin G(p^*,q^*)$. □

II.34. COROLLARY. The class of strongly conclusive connectors is not characterizable.

II.35. COROLLARY. The logic determined by the class of selection functions is $\mathbf{B}$.

II.2.2. Constraints on Antecedents

The next definition mirrors definition 22.

II.36. DEFINITION. Let F be a connector with domain W.
F is *inclusive* iff for every $i \in W$ and $p,q,r \subset W$,
(i) $i \in F(p,p)$;
(ii) if $i \in F(q,p)$ and $r \subset q$, then $i \in F(r,p)$;
(iii) if $i \in F(q,p)$ and $i \in F(r,p)$, then $i \in F(q \cup r, p)$.
F is *strongly inclusive* if (iii) can be strengthened to
(iii)' if $Q \neq \emptyset$ and $i \in F(q,p)$ for every $q \in Q$, then $i \in F(\cup Q,p)$. □

Let F be any connector. Let A be the antecedence function derived from F. Saying that F is inclusive amounts to saying that for every world i and proposition p, the set $A_i(p)$ - i.e. the set of antecedents of p at i - constitutes an *ideal* (that contains p as an element): every r that is included by the union $q_1 \cup \ldots \cup q_n$ of finitely many antecedents $q_1,\ldots,q_n$ at i is again an antecedent of p at i. Saying that F is strongly inclusive amounts to saying that each $A_i(p)$ constitutes a *principal* ideal.

In this section we will take a closer look at connectors that are both conclusive and inclusive. These are interesting for several reasons, one being that the logic determined by the class of connectors is the weakest conditional logic in which the arrow $\rightarrow$ behaves as a *strict* implication. The following will make this clear.

II.37. PROPOSITION. Let F be conclusive and inclusive.

(i) $F(p,q) = F(p \cap (W \sim q), \emptyset) = F(W, (W \sim p) \cup q)$.

(ii) F is strongly conclusive iff F is strongly inclusive.

PROOF.

(i) The following scheme will suffice.

$$i \in F(p,q) \overset{i}{\Rightarrow} \left.\begin{array}{l} i \in F(p \cap (W \sim q), q) \\ i \in F(p \cap (W \sim q), (W \sim q)) \end{array}\right\} \overset{c}{\Rightarrow} i \in F(p \cap (W \sim q), \emptyset)$$

$$i \in F(p \cap (W \sim q), \emptyset) \overset{c}{\Rightarrow} \left.\begin{array}{l} i \in F(p \cap (W \sim q), (W \sim p) \cup q) \\ i \in F(W \sim p) \cup q, (W \sim p) \cup q) \end{array}\right\} \overset{i}{\Rightarrow} i \in F(W, (W \sim p) \cup q)$$

$$i \in F(W, (W \sim p) \cup q) \overset{i}{\Rightarrow} \left.\begin{array}{l} i \in F(p, (W \sim p) \cup q) \\ i \in F(p, p \cup q) \end{array}\right\} \overset{c}{\Rightarrow} i \in F(p,q)$$

(Here 'i' means 'by inclusiveness', and 'c' means 'by conclusiveness'.)

(ii) Suppose F is inclusive and strongly conclusive. We show that F is strongly inclusive. Assume that $Q \neq \emptyset$ and that $i \in F(q,p)$ for every $q \in Q$. By (i) it follows that $i \in F(W, (W \sim q) \cup p)$ for every $q \in Q$. Since F is strongly conclusive, and $\cap\{r \mid r = (W \sim q) \cup p \text{ for some } q \in Q\} = (W \sim \cup Q) \cup p$, this means that $i \in F(W, (W \sim \cup Q) \cup p)$. Applying (i), we find that $i \in F(\cup Q, p)$, which is what we had to show.

The proof of the converse is the dual of this. □

Now, let '$\bot$' be an abbreviation of '$(\phi_0 \wedge \sim\phi_0)$', where ϕ_0 is some fixed atomic sentence; and let '$\Box\phi$' be short for '$(\sim\phi \rightarrow \bot)$'. Using these abbreviations, we find:

II.38. PROPOSITION. Let K be the class of conclusive and inclusive connectors. The following arguments are valid within K.

(i) $\phi \rightarrow \psi / \Box(\sim\phi \vee \psi)$;

(ii) $\Box(\sim\phi \vee \psi) / \phi \rightarrow \psi$;

(iii) $\Box(\phi \wedge \psi) / \Box\phi$;

(iv) $\Box\phi \wedge \Box\psi / \Box(\phi \wedge \psi)$;

(v) $\emptyset / \Box(\phi \vee \sim\phi)$.

PROOF. (i) and (ii) are immediate consequences of proposition 37. The proofs of (iii), (iv) and (v) are straightforward. □

This confirms what was said above: within the class of conclusive and inclusive connectors, the arrow $\rightarrow$ behaves like a strict implication: (i) and (ii) say that within this class $\phi \rightarrow \psi$ is equivalent to $\Box(\sim\phi \vee \psi)$, while (iii), (iv) and (v) add that $\Box$ is the necessity operator of some normal modal logic.

The question is *which* normal modal logic.

II.39. DEFINITION. Let W be a set of possible worlds. An *accessibility-function* for W is any function D that assigns to each world i in W a subset D_i of W.
If $j \in D_i$, we say that j is *accessible from* i. □

II.40. DEFINITION. Let [ ] be a standard interpretation over a given set W of worlds, and let D be an accessibility function for W. [ ] is *based on* D iff for any two sentences ϕ and ψ,

$$i \in [\phi \rightarrow \psi] \text{ iff } [\phi] \cap D_i \subset [\psi].$$ □

What these definitions tell is that a conditional $\phi \rightarrow \psi$ is true at a given world i just in case the consequent ψ is true at all $[\phi]$-worlds that are in some sense accessible from i.[12)] It may seem that this does not add very much to definitions 26 and 27 where selection functions were introduced - and, of course, as long as we do not precisely know what 'accessible' means the above cannot count as a theory any more than those definitions did. Still, formally, the transition from 'all $[\phi]$-worlds that are relevant to $[\phi]$ at i' to 'all $[\phi]$-worlds that are accessible from i' parallels the one from strongly conclusive connectors to connectors that are both strongly conclusive and inclusive.

(I suppose you have recognized Kripke's semantics for normal modal logics in the above. If not, notice that according to definitions 39 and 40 a sentence of the form

$\Box\phi$ is true at a given world i just in case ϕ is true at all worlds accessible from i.)

II.41. DEFINITION. Let F be a connector with domain W, and $\mathcal{D}$ an accessibility function for W. F is *interchangeable* with $\mathcal{D}$ iff for every $i \in W$ and $p,q \subset W$,

$$i \in F(p,q) \text{ iff } p \cap \mathcal{D}_i \subset q. \qquad \Box$$

II.42. THEOREM. A connector F is interchangeable with an accessibility function iff F is strongly conclusive and inclusive.

PROOF. It is left to the reader to check that F is strongly conclusive and inclusive if F is interchangeable with some accessibility function $\mathcal{D}$.

To prove the converse, suppose that F is strongly conclusive and inclusive. Let $\mathcal{D}$ for every i in the domain W of F be defined by

$$\mathcal{D}_i = \cap C_i(W),$$

where C is the consequence function derived from F.
We must show that

$$i \in F(p,q) \text{ iff } p \cap \mathcal{D}_i \subset q.$$

Suppose $i \in F(p,q)$. Then by proposition 37, $i \in F(W,(W \sim p) \cup q)$. In other words, $(W \sim p) \cup q \in C_i(W)$. Therefore, $p \cap \mathcal{D}_i = p \cap \cap C_i(W) \subset q$. Conversely, suppose $p \cap \mathcal{D}_i \subset q$. Then $\cap C_i(W) \subset (W \sim p) \cup q$. Given that F is strongly conclusive, this means that $(W \sim p) \cup q \in C_i(p)$. In other words, $i \in F(W,(W \sim p) \cup q)$. By proposition 37, this means that $i \in F(p,q)$. $\Box$

II.43. THEOREM. Let F be both conclusive and inclusive. There is an ultrafilter extension G of F such that G is strongly conclusive and inclusive.

PROOF. We follow the line of proof of lemma 33.
Suppose W is the domain of F, and let U and * be defined as in definition 30. Consider for every $u \in U$, the set

$$\delta_u = \{r \mid F(W,r) \in u\}.$$

(i) δ_u is a filter on $\mathcal{W}$. (Compare the proof of theorem 33.)

(ii) $F(p,q) \in u$ iff $(\mathcal{W} \sim p) \cup q \in \delta_u$. (Apply proposition 37.)

Now consider the accessibility function $\mathcal{D}$ for $\mathcal{U}$ defined by

$$\mathcal{D}_u = \{v \in \mathcal{U} \mid \delta_u \subset v\}.$$

Let G be the connector interchangeable with $\mathcal{D}$, i.e. let G for every P, Q $\subset$ u be given by

$$u \in G(P,Q) \text{ iff } P \cap \mathcal{D}_u \subset Q.$$

By theorem 33, G is strongly conclusive and inclusive. It remains to show that G is an ultrafilter extension of F. Suppose $u \notin (F(p,q))^*$. This is so iff $F(p,q) \notin u$. Given (ii) this means that $(\mathcal{W} \sim p) \cup q \notin \delta_u$. Given (i), this means there is some ultrafilter v such that $\delta_u \subset v$ and $p \cap (\mathcal{W} \sim q) \in v$. This is in turn so in case there is some $v \in \mathcal{D}_u$ such that $v \in p^*$ and $v \notin q^*$. Which by the interchangeability of $\mathcal{D}$ and G amounts to $u \notin G(p^*,q^*)$. □

It follows as a corollary that the logic determined by the class of all conclusive and inclusive connectors and the logic determined by the class of all strongly conclusive and inclusive connectors are the same. As a corollary of theorem 42 we have that the latter is equivalent to the weakest normal modal logic, better known as $\mathbb{K}$.

Here is a suitable description of this logic.

II.44. THEOREM. Let $\mathbb{K}$ be the weakest logic containing all arguments of the form CI, CW, CC, and
Strengthening the Antecedent (AS): $\phi \rightarrow \psi / (\phi \wedge \chi) \rightarrow \psi$.
Disjunction of Antecedents (AD) : $\phi \rightarrow \psi, \chi \rightarrow \psi / (\phi \vee \chi) \rightarrow \psi$.
$\mathbb{K}$ is complete and its range $K_{\mathbb{K}}$ is the class of all conclusive and inclusive connectors.

PROOF. Obviously, $F \in K_{\mathbb{K}}$ iff F is conclusive and inclusive. To prove the completeness of $\mathbb{K}$, it is sufficient to prove that $\mathbb{K}$ is canonical.
Let [ ] be any standard, compositional intepretation verifying $\mathbb{K}$. Let $\mathcal{W}$ be the set of worlds over which [ ] is defined. Now consider the connector F with domain $\mathcal{W}$ given by

$i \in \mathcal{F}(p,q)$ iff there is a sentence χ such that $i \in [\Box\chi]$ and $[\chi] \subset (\mathcal{W} \sim p) \cup q$.
It is left to the reader to check that [ ] can be based on $\mathcal{F}$, and that $\mathcal{F} \in \mathcal{K}_{\mathbb{K}}$. The proof is straightforward and makes use of the following lemma.

II.45. LEMMA. The arguments mentioned in proposition 38 all belong to $\mathbb{K}$.

PROOF. I will only prove (i). In a quasi-derivation this proof looks like this. (Applications of Replacement are left out.)

$$\dfrac{\dfrac{\phi \to \psi}{(\phi \wedge \sim\psi) \to \psi}\ (AS) \qquad \dfrac{\dfrac{\emptyset}{(\phi \wedge \sim\psi) \to (\phi \wedge \sim\psi)}\ (CI)}{(\phi \wedge \sim\psi) \to \sim\psi}\ (CW)\ (CW)}{(\phi \wedge \sim\psi) \to \bot}\ (CC)$$

□

In working out the proof of theorem 44, you will have noticed that $\mathbb{K}$ can alternatively be described as the weakest logic containing all arguments mentioned in the lemma. Here is another description of $\mathbb{K}$:
$\mathbb{K}$ is the weakest conditional logic containing all arguments of the form $\phi \to \psi / \sim\psi \to \sim\phi$ (Contraposition).
This underlines once more that inclusiveness is just the dual of conclusiveness. Yet another description of $\mathbb{K}$ is this:
$\mathbb{K}$ is the weakest logic containing all arguments of the form

- $\emptyset / \phi \to (\phi \vee \psi)$ (Logical Implication);
- $\phi \to \psi / \phi \to (\phi \wedge \psi)$ (Minimal Conjunction);
- $\phi \to \psi / (\phi \vee \psi) \to \psi$ (Minimal Disjunction);
- $\phi \to \psi, \psi \to \chi / \phi \to \chi$ (Hypothetical Syllogism).

I mention this description mainly because it shows how strong the Hypothetical Syllogism is. Logical Implication, Minimal Conjunction and Minimal Disjunction are very weak, even all together.

The majority of possible-worlds semanticists is convinced that the logic of natural language conditionals cannot be equated with $\mathbb{K}$, or with any of its extensions. More precisely, most possible-worlds semanticists are convinced that counterfactuals do not behave as strict implications, and many would add that indicatives do not do so either.

The scapegoat here is the principle of Strengthening the Antecedent, but, of course, where this principle is rejected, stronger principles (stronger when added to $\mathbb{B}$, that is) like Contraposition and the Hypothetical Syllogism will be rejected too.

Intuitive counterexamples to the principle of Strengthening the Antecedent exist in abundance. As for counterfactuals, one might even go as far as saying that it is difficult to think up a contingent premise of the form

if it had been the case that ϕ, *then it would have been the case that* ψ

without there being a sentence χ ready at hand to render

if it had been the case that ϕ *and* χ, *then it would have been the case that* ψ

an absurd conclusion.

Yet, on the other hand, it is no less difficult to think up a premise of the form

if it had been the case that ϕ *or* χ, *then it would have been the case that* ψ

from which it *is* absurd to conclude

if it had been the case that ϕ, *then it would have been the case that* ψ.

Unfortunately, within the framework of possible-worlds semantics, it is impossible to reject $\phi \rightarrow \psi/(\phi \wedge \chi) \rightarrow \psi$ while retaining $(\phi \vee \chi) \rightarrow \psi/\phi \rightarrow \psi$. These two argument forms characterize the same class of connectors.

What now? Give up possible-worlds semantics, and try to develop a theory in which $\phi \rightarrow \psi/(\phi \wedge \chi) \rightarrow \psi$ is not valid but $(\phi \vee \chi) \rightarrow \psi/\phi \rightarrow \psi$ is? Of course not. As we saw in section I.2.3, where this problem was discussed in a more general setting, the ultimate criterion is this: which *combination* of syntax + semantics + pragmatics gives rise to the most convincing explanations? So, even if we leave syntax in peace (as agreed), there still are several ways out. We could try to explain away the intuitively sound instances of $(\phi \vee \chi) \to \psi / \phi \to \psi$ as being *pragmatically correct* rather than logically valid. Or, alternatively, we could try to explain away the intuitively absurd instances of $(\phi \to \psi)/(\phi \wedge \chi) \to \psi$ as being *pragmatically incorrect* rather than logically invalid.

We are not yet in a position to decide which option is best. Still, it is worthwhile looking at the consequences of rejecting the principle of Strengthening the Antecedent for our formal set-up. Clearly, an immediate consequence will be that we must give up the idea that our eventual truth definition can take the simple form it took in definition 40:

$i \in [\phi \to \psi]$ iff for all worlds $j \in [\phi] \cap \mathcal{D}_i$ it holds that $j \in [\psi]$.

Now, it might be hoped - or even expected - that at least part of this definition can be saved. Wouldn't it be nice if we could just replace the quantifier *all* by a less inclusive one? For example, what about

$$i \in [\phi \to \psi] \text{ iff for } \left.\begin{matrix} \textit{most} \\ \textit{many} \\ \textit{some} \end{matrix}\right\} \begin{matrix} \text{worlds } j \in [\phi] \cap \mathcal{D}_i \\ \text{it holds that } j \in [\psi]? \end{matrix}$$

Unfortunately, this cannot be done. There are lots of quantifiers that give the arrow $\to$ the properties of a nonstrict implication, but on finite domains all of them boil down to *all*. This means that they do refute the principle of Strengthening the Antecedent, but only so for infinite sets of possible worlds.

Clearly, that is not what we are looking for. If we reject the principle of Strengthening the Antecedent, then certainly our reasons for doing so will have nothing to do with the cardinality of the set of possible worlds. (Or would you say that the absurdity of the argument

If there had been sugar in this coffee, it would have tasted better

---

∴ *If there had been sugar and dieseloil in this coffee, it would have tasted better*

is ultimately due to the fact that an infinite, rather than a finite number of alternatives is involved?)

Let us turn to the proof of the mentioned result. I have extracted it from Van Benthem (1984), a paper in which conditionals are studied from the perspective of the theory of generalized quantifiers.

To understand the proof, there are a few things about quantifiers you should know.

A quantifier can best be understood as a functor $\mathcal{Q}$ that assigns to any domain W a relation between the subsets of W.

Thus, for a set W of possible worlds, and a pair p,q of propositions, '$\mathcal{Q}_W(p,q)$' will mean that the relation $\mathcal{Q}$ holds in W between p and q. (Read '$\mathcal{Q}_W(p,q)$' as 'within W, $\mathcal{Q}$ p-worlds are q-worlds'.) In this fashion, the familiar quantifiers *all*, *some*, and *most* can be introduced as follows:

$all_W(p,q)$ iff $|p \sim q| = 0$;

$some_W(p,q)$ iff $|p \cap q| \neq 0$;

$most_W(p,q)$ iff $|p \cap q| > |p \sim q|$.

(Here, '$|p|$' denotes the cardinal number of the set p.)

A less familiar quantifier is the quantifier M defined by

$$M_W(p,q) \quad \text{iff} \quad \begin{cases} |p \sim q| = 0, & \text{when p is finite} \\ |p \sim q| < |p|, & \text{when p is infinite.} \end{cases}$$

We will read '$M_W(p,q)$' as '*by far the most* p-worlds are q-worlds', keeping in mind that for finite p '*by far the most* p-worlds' means '*all* p-worlds'.

Of course, that a functor assigns a relation between its subsets to any set is not enough for it to count as a quantifier. In fact, there are quite a number of conditions that such a functor must satisfy. Here only one condition matters.

Quantity: $\mathcal{Q}$ is a quantifier only if for any W and W', any bijection σ from W onto W', and any $p, q \subset W$,

$$\mathcal{Q}_W(p,q) \text{ iff } \mathcal{Q}_{W'}(\sigma[p],\sigma[q]).$$

(Here $\sigma[p] = \{j \in W' \mid j = \sigma(i) \text{ for some } i \in p\}$.)

What this condition says is that if the relation $\mathcal{Q}_W$ holds between p and q, and you systematically replace the elements of W by other ones, then the relation $\mathcal{Q}_{W'}$ must also hold between the new sets p' and q'. A quantifier $\mathcal{Q}$ is, so to speak, indifferent to the specific nature of the elements in its domain, the only thing that matters is their contribution to the *size* of the sets containing them.

Apart from general principles, like Quantity, there are also special principles that hold for certain kinds of quantifiers, but not for all of them. One such principle is the following.

Activity: if $p \neq \emptyset$ then there are q and q' such that $\mathcal{Q}_W(p,q)$ and not $\mathcal{Q}_W(p,q')$.

What this means is this: the question whether $\mathcal{Q}_W(p,q)$ or not $\mathcal{Q}_W(p,q)$ may not depend only on p, except perhaps when $p = \emptyset$; somewhere q must come in.

Note that Activity does not hold for everything that one might want to call a quantifier. Take the quantifier *less than two*, for example, which is given by

less than two$_W$ (p,q) iff $|p \cap q| < 2$.

Clearly, if $|p| < 2$ then less than two p-worlds are q-worlds, no matter how many q-worlds there are.

Recall, however, that we are interested in quantifiers that can be used in a truth definition for conditionals. We want to equate '$i \in [\phi \rightarrow \psi]$' with '$\mathcal{Q}_W([\phi] \cap \mathcal{D}_i,[\psi])$'. And we certainly do no want a sentence of the form $\phi \rightarrow \psi$ ever to be true (or false), at a given world i merely because the set $[\phi] \cap \mathcal{D}_i$ contains a particular number of elements (except perhaps when this particular number is zero). The number of $[\psi]$-worlds, in particular the number of $[\psi]$-worlds belonging to $[\phi] \cap \mathcal{D}_i$, should always remain relevant.

II.46. PROPOSITION. Let $\mathcal{Q}$ be any active quantifier. The first of the following statements implies the second.

(i) For any W, and $p,q,r \subset W$,

(a) $\mathcal{Q}_W(p,p)$;

(b) if $\mathcal{Q}_W(p,q)$ and $q \subset r$, then $\mathcal{Q}_W(p,r)$;

(c) if $\mathcal{Q}_W(p,q)$ and $\mathcal{Q}_W(p,r)$, then $\mathcal{Q}_W(p,q \cap r)$.

(ii) For any W, and $p,q \subset W$,

(d) if $all_W(p,q)$ then $\mathcal{Q}_W(p,q)$;

(e) if $\mathcal{Q}_W(p,q)$ then *by far the most*$_W$ (p,q).

PROOF. It is left to the reader to check that (i) implies (d).

Now assume that (i) and suppose that (e) does not hold. Then we can find a set W, and $p,q \subset W$ such that $\mathcal{Q}_W(p,q)$, while either (*) p is finite but $|p \sim q| \neq 0$

or (**) p is infinite but $|p \sim q| = |p|$.

Claim: in both these cases it follows that $\mathcal{Q}_W(p,\emptyset)$.

For, suppose (*). Note first that by (a) and (c) we have that $\mathcal{Q}_W(p,p \cap q)$. Let j be any world in $p \cap q$, and let k be any world in $p \sim q$. Consider the bijection σ from W onto W such that $\sigma(j) = k$ and $\sigma(k) = j$ while for all i different from j and k, $\sigma(i) = i$. By Quantity, $\mathcal{Q}_W(\sigma[p],\sigma[p \cap q])$, which means that $\mathcal{Q}_W(p,((p \cap q) \sim \{j\}) \cup \{k\})$. Given that $\mathcal{Q}_W(p,p \cap q)$, it follows by (c) that $\mathcal{Q}_W(p,(p \cap q) \sim \{j\})$. Since this holds for any $j \in p \cap q$, and since $p \cap q$ is finite we find that $\mathcal{Q}_W(p,\emptyset)$. (Apply (c).)

Now suppose (**). Again we have that $\mathcal{Q}_W(p,p \cap q)$. Since $|p \sim q| = |p|$ and $|p \cap q| \leq |p|$ we can find a set $r \subset p \sim q$ such that $|r| = |p \cap q|$. Consider any bijection from W onto W such that $\sigma[p] = p$ and $\sigma[p \cap q] = r$. Applying Quantity, we find that $\mathcal{Q}_W(p,r)$. From this and the fact that $\mathcal{Q}_W(p,p \cap q)$ it follows by (c) that $\mathcal{Q}_W(p,\emptyset)$.

This proves the claim.

Given that $\mathcal{Q}_W(p,\emptyset)$ it follows by (b) that $\mathcal{Q}_W(p,r)$ for all r. By Activity this means that $p = \emptyset$. But this contradicts both (*) and (**). □

II.47. COROLLARY. An active quantifier satisfies the basic conditional logic $\mathcal{B}$ only if it is at least as strong as *by far the most*, and not stronger than *all*.

PROOF. If we replace the last line of definition 40 by

$$i \in [\phi \rightarrow \psi] \text{ iff } \mathcal{Q}_{\mathcal{W}}([\phi] \cap \mathcal{D}_i, [\psi]),$$

then this will yield a logic at least as strong as $\mathcal{B}$ iff the quantifier $\mathcal{Q}$ concerned satisfies for any $\mathcal{W}$, any $\mathcal{D}$, any i, and any p, q, r the following conditions.

(a)' $\mathcal{Q}_{\mathcal{W}}(p \cap \mathcal{D}_i, p)$

(b)' if $\mathcal{Q}_{\mathcal{W}}(p \cap \mathcal{D}_i, q)$ and $q \subset r$, then $\mathcal{Q}_{\mathcal{W}}(p \cap \mathcal{D}_i, r)$

(c)' if $\mathcal{Q}_{\mathcal{W}}(p \cap \mathcal{D}_i, q)$ and $\mathcal{Q}_{\mathcal{W}}(p \cap \mathcal{D}_i, r)$,
then $\mathcal{Q}_{\mathcal{W}}(p \cap \mathcal{D}_i, q \cap r)$

It is not difficult to see that this is so just in case (i) of proposition 46 holds. Since (i) implies (ii), the corollary follows immediately. □

Finally, note that there are quantifiers between *by far the most* and *all* for which the principle of Strengthening the Antecedent does not hold. In fact we have that for any $\mathcal{W}$ such that $|\mathcal{W}| \geq \omega$ and any $p,q,r \subset \mathcal{W}$ such that $r \subset p \subset q$, $|r| < \omega$, and $|p| = |q| = \omega$ that *by far the most*$_{\mathcal{W}}(p,q)$ but not *by far the most*$_{\mathcal{W}}(p \cap r, q)$. However, if $|\mathcal{W}| < \omega$ such counterexamples do not exist. If $|\mathcal{W}| < \omega$, *by far the most*$_{\mathcal{W}}(p,q)$ just in case *all*$_{\mathcal{W}}(p,q)$.[13)]

II.3. ALTERNATIVES

II.3.1. Ramsey's suggestion and Stalnaker's theory

The logical problem of conditionals was first introduced into possible-worlds semantics by Robert Stalnaker in his by now classic 1968 paper 'A Theory of Conditionals'. Taking the explication of logical validity in terms of truth for granted, Stalnaker saw the logical problem of conditionals as the problem of giving their *truth* conditions, a problem that according to him must be well distinguished from the problem of giving their *assertability* conditions. On his view the question as to when one can justifiably believe or assert a conditional sentence belongs to pragmatics, and not to semantics.

Still, it is an in this sense pragmatic view which Stalnaker takes as a starting point in solving his logical problem: the test for evaluating a conditional statement suggested by Frank Ramsey in 1929 and which is summed up by Stalnaker (op. cit.: 102) as follows:

> *This is how to evaluate a conditional: first, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is then true.*

Ramsey's original suggestion only covered the case in which the antecedent is consistent with 'your' stock of beliefs. In that case, which is typical of indicative conditionals, no adjustments are required. Following an idea of Rescher (1964), Stalnaker generalizes this to the case in which the antecedent cannot simply be added to 'your' stock of beliefs without introducing a contradiction. In this case, which is typical of counterfactuals, adjustments are required.

Now, in effect what Stalnaker does is to reconstruct the above belief conditions as truth conditions. He thinks that truth may not be allowed to depend on beliefs, that you have to appeal to the facts. So in his rebuilt version the actual world plays the role of your stock of beliefs in Ramsey's. Adding the antecedent and making the necessary changes becomes, in Stalnaker's version, moving to that possible world at which the antecedent is true and which in all other respects most resembles the actual world. Here is a first approximation of the truth conditions Stalnaker (op. cit.: 102) has in mind.

Consider the possible world in which ϕ is true, and which otherwise differs minimally from the actual world. 'If ϕ then ψ' is true (false) just in case ψ is true (false) in that world.

The obvious question arises as to which world if any most resembles the actual one in the other respects.

This is a slippery question. According to Stalnaker it is also in essence a pragmatic question and has nothing to do with his logical problem. He stresses the point (op. cit.: 109) that '*the context of utterance, the purpose of the assertion, and the beliefs of the speaker or his community*' may make a difference to the particular world which has the property concerned. *How* these contextual features make that difference is less important. The only thing that matters is *that* these differences can be made.

The groundwork for a formal elaboration of Stalnaker's ideas is provided by definition 26 and 27 where selection functions are introduced and where it is explained how these can serve as a basis for the interpretation of conditionals. What Stalnaker adds to this formally amounts to

II.48. DEFINITION. Let S be a selection function for W. S is *stalnakerlike* iff for every $i \in W$ and $p, q \subset W$ the following holds.

(i) Uniqueness : if $S_i(p) \neq \emptyset$ then $S_i(p) = \{j\}$ for some $j \in W$;

(ii) Centering : if $i \in p$ then $S_i(p) = \{i\}$;

(iii) Equivalence: if $S_i(p) \subset q$ and $S_i(q) \subset p$ then $S_i(p) = S_i(q)$. □

Let S be a selection function for W. Consider $S_i(p)$ for arbitrary i and p. We have learned to think of the worlds $i \in S_i(p)$ as the p-worlds that are - somehow - *relevant* to p at i. Now, (i)-(iii) all go without saying if you believe, as Stalnaker does, that only the p-world most similar to i can be relevant to p at i. (i) says that at most one p-world really matters. With the aid of (ii), we can partially answer the question which one: if p holds at i, then i itself is the world relevant to p at i. (Of course: if p holds at i then i itself is the p-world most similar to i.) And (iii) provides an identity criterion for the case that p does not hold at i: if q holds at the world relevant to p at i and p holds at the world relevant to q at i, then these worlds are identical. (Of course if q holds at the p-world most resembling i, then this world will also be the $p \cap q$-world most resembling i, because any $p \cap q$-world more resembling i would also be a p-world more resembling i. Likewise if p holds at the q-world most resembling i, this q-world, too, must be the $p \cap q$-world most resembling i.)

In the informal presentation of Stalnaker's ideas I did not prepare the reader for the fact that $S_i(p)$ will sometimes be the empty set. If $p = \emptyset$, for example, there is no p-world, let alone a p-world most similar to i. Stalnaker allows that $S_i(p) = \emptyset$ even in cases where $p \neq \emptyset$ - the underlying idea being that in some cases the worlds at which p holds might all be so different from i that it doesn't make sense to ask for the one most similar to i. In such a case, it makes no sense to suppose that p, not even counterfactually. Consequently, any conditional whose antecedent expresses p will be true; *ex absurdo sequitur quodlibet*.

If you disagree with Stalnaker here - believing, as you may, that it is never absurd to suppose a proposition that is not itself absurd - then you may want to constrain the selection functions more than is done by Stalnaker. You may want to add the requirements that S is universal in the following sense.

Universality: for every $p \subset W$, $S_i(p) = \emptyset$ only if $p = \emptyset$.

You may also wonder whether perhaps the selection functions have not been constrained too much already. Here the Uniqueness Assumption laid down in (i) of definition 48 is, I think, a case in point. Can we really be sure that there will always be, for any proposition p and world i, at most one p-world most resembling i? Couldn't there be several such p-worlds, all equally close to i and all closer to i than any other world - all minimally differing from i in respects other than p, but each differing from i in other respects?

II.49. EXAMPLE. Consider the six marble worlds as depicted in table 1 of section I.1.1. Which of these is the world at which the blue marble is in box 2 and which otherwise differs minimally from world I? □

The example serves to set a problem rather than to solve it. It is terribly difficult to choose between the worlds II, III and IV here, but it would be premature to conclude that it is impossible. The point is that Stalnaker's theory does not tell us how to choose. That is no logical question, Stalnaker says, and he delegates it to the pragmatics. But in the absence of the answer, the Uniqueness Requirement remains a mere stipulation with a questionable *logical* pay off.

The logical pay off of the Uniqueness Requirement is the principle of the Conditional Excluded Middle

$(\phi \to \psi) \vee (\phi \to \sim \psi)$ (see below). Hence at least one of the following sentences is true at marble-world I:

If the blue marble had been in box 2, the red marble would have been in box 1.

If the blue marble had been in box 2, the red marble would not have been in box 1.

Quite a few people happen to think that both these sentences are false at world I. They could be wrong. According to Stalnaker they are wrong. But his theory does not offer any explanation as to why.

Stalnaker speaks freely of the p-world most similar to i. If the above criticism is correct one should rather speak of *the set* of p-worlds most similar to i. But perhaps even that cannot always be done. Perhaps there are propositions p such that for every p-world j a world exists that is more similar to i than j is. So that one can get closer and closer to i without ever getting in a p-world that is closest to i. Lewis (1973: 20) argues that this might very well occur: "Suppose we entertain the counterfactual supposition that at this point

---

there appears a line more than an inch long. (Actually it is just under an inch.) There are worlds with a line 2" long; worlds presumably closer to ours with a line 1½" long; worlds presumably still closer to ours with a line 1¼" long; worlds presumably still closer ... But how long is the line in the *closest* worlds with a line more than an inch long? If it is 1 + x" for any x however small, why are there not other worlds still closer to ours in which it is 1 + ½x", a length still closer to its actual length? The shorter we make the line (above 1"), the closer we come to the actual length; so the closer we come, presumably, to our actual world. Just as there is no shortest possible length above 1", so there is no closest world to ours among the worlds with lines more than an inch long ..."

In the next section we will discuss the modifications of Stalnaker's theory which Lewis is calling for.

On to the logic yielded by Stalnaker's theory. How does it relate to the logics discussed in the previous chapter? Perhaps the quickest way to answer this is to see what properties a connector F must have in order to be interchangeable with a stalnakerlike selection function.

From theorem 29, we already know that F must be strongly conclusive: for any i and p, the set

$$C_i(p) = \{q \mid i \in F(p,q)\} = \{q \mid S_i(p) \subset q\} \quad (*)$$

must constitute a principal filter that contains p as an element. Given theorem 35, this means that Stalnaker's logic will be at least as strong as $\mathcal{B}$.

What do the conditions (i), (ii) and (iii) of definition 48 add to this? The Uniqueness Condition says that for any i and p, $C_i(p)$ must be a *maximal* filter. (More precisely, $C_i(p)$ must be either the improper filter or an ultrafilter generated by $\{j\}$ for some world j.) This means that F, in addition to being strongly conclusive, must have the property that for any i, p, and q,

$$\text{either } i \in F(p,q) \text{ or } i \in F(p, W \sim q).$$

Obviously, the class of connectors with this property is characterized by the principle of

Conditional Excluded Middle (CEM): $\emptyset/(\phi \to \psi) \vee (\phi \to \sim\psi)$.

The centering condition can be split up into

(i) if $i \in p$ then $i \in S_i(p)$; and

(ii) if $i \in p$ then $S_i(p) \subset \{i\}$.

By making use of (*), we see that (i) can be restated for F as

if $i \in p$, then $i \in q$ for every q such that $i \in F(p,q)$.

Clearly, the corresponding argument form is

Modus Ponens (MP): $\phi, \phi \to \psi/\psi$.[14]

Notice that (ii) is already covered by Uniqueness. In terms of F, it runs like this:

if $i \in p$ and $i \in q$ then $i \in F(p,q)$.

So, the corresponding argument form is

Conjunctive Sufficiency (CS): $\phi \wedge \psi/\phi \to \psi$.

Finally, the constraint that the Equivalence Condition imposes on F is this:

if $i \in F(p,q)$, $i \in F(q,p)$ and $i \in F(p,r)$, then $i \in F(q,r)$,

which gives rise to the following characteristic argument form

Conditional Equivalence (CE): $\phi \to \psi, \psi \to \phi, \phi \to \chi / \psi \to \chi$.

II.50. THEOREM. A connector F is interchangeable with a stalnakerlike selection function iff F is strongly conclusive, and for any world i and propositions p, q, r in the domain of F,

(i) either $q \in F(i,p)$ or $(W \sim q) \in F(i,p)$;

(ii) if $i \in p$ and $i \in F(p,q)$, then $i \in q$;

(iii) if $i \in F(p,q)$, $i \in F(q,p)$, and $i \in F(p,r)$, then $i \in F(q,r)$.

PROOF. The proposition sums up the preceding discussion. □

The next proposition must be compared with definition 36. It says among other things how inclusive a connector is if it is interchangeable with a stalnakerlike selection function.

II.51. PROPOSITION. In theorem 50, condition (iii) can be replaced by

(iv) if $i \in F(q,p)$ then $i \in F(q \cap r,p)$ *provided that* $i \in F(q,r)$;

(v) if $Q \neq \emptyset$ and $i \in F(q,p)$ for every $q \in Q$, then $i \in F(\cup Q,p)$.

PROOF. Let (c) be short for 'F is conclusive'. First we show that (iv) follows from (c) + (iii).

$$\left.\begin{array}{l}\left.\begin{array}{l}\boxed{i \in F(q,r)} \\ \overset{c}{\Rightarrow}\; i \in F(q,q)\end{array}\right\} \overset{c}{\Rightarrow}\; i \in F(q,r \cap q) \\ \qquad\qquad\qquad\;\; \overset{c}{\Rightarrow}\; i \in F(r \cap q,q) \\ \qquad\qquad\qquad\quad\;\; \boxed{i \in F(q,p)}\end{array}\right\} \overset{iii}{\Rightarrow}\; i \in F(q \cap r,p)$$

(v) follows from (ii), (iii) and the assumption that F is strongly conclusive.

Note that according to (i) for any $q \in \mathcal{Q}$, either $i \in \mathcal{F}(\bigcup\mathcal{Q},q)$ or $i \in \mathcal{F}(\mathcal{Q},\mathcal{W} \sim q)$. So then there are two cases:
First, assume that there is some $q \in \mathcal{Q}$, such that $i \in \mathcal{F}(\bigcup\mathcal{Q},q)$.
Then the proof can be completed this way:

$$\left.\begin{array}{l} \quad i \in \mathcal{F}(\bigcup\mathcal{Q},q) \\ \overset{c}{\Rightarrow}\ i \in \mathcal{F}(q,\bigcup\mathcal{Q}) \\ \quad \boxed{i \in \mathcal{F}(q,p)} \end{array}\right\} \overset{iii}{\Rightarrow} i \in \mathcal{F}(\bigcup\mathcal{Q},p)$$

Now, assume that for every $q \in \mathcal{Q}$, $i \in \mathcal{F}(\bigcup\mathcal{Q},\mathcal{W} \sim q)$. Since $\mathcal{F}$ is strongly conclusive and $\bigcap\{\mathcal{W} \sim q \mid q \in \mathcal{Q}\} = \mathcal{W} \sim \bigcup\mathcal{Q}$ we then have

$$\left.\begin{array}{l} \quad i \in \mathcal{F}(\bigcup\mathcal{Q},\mathcal{W} \sim \bigcup\mathcal{Q}) \\ \overset{c}{\Rightarrow}\ i \in \mathcal{F}(\bigcup\mathcal{Q},\bigcup\mathcal{Q}) \end{array}\right\} \overset{c}{\Rightarrow} i \in \mathcal{F}(\bigcup\mathcal{Q},\emptyset) \overset{c}{\Rightarrow} i \in \mathcal{F}(\bigcup\mathcal{Q},p)$$

Finally we show that (iii) is implied by (c) + (iv) + (v).

$$\boxed{i \in \mathcal{F}(p,r)}\ \left.\begin{array}{l} \overset{c}{\Rightarrow} i \in \mathcal{F}(p,(\mathcal{W}\sim p) \cup r) \\ \overset{c}{\Rightarrow} i \in \mathcal{F}((\mathcal{W}\sim p) \cap q,(\mathcal{W}\sim p) \cup r) \end{array}\right\} \overset{v}{\Rightarrow} i \in \mathcal{F}(p \cup q,(\mathcal{W}\sim p) \cup r) \quad (*)$$

$$\left.\begin{array}{l} \left.\begin{array}{l} \quad \boxed{i \in \mathcal{F}(q,p)} \\ \overset{c}{\Rightarrow}\ i \in \mathcal{F}(p,p) \end{array}\right\} \overset{v}{\Rightarrow} i \in \mathcal{F}(p \cup q,p) \\ \overset{*}{\Rightarrow} i \in \mathcal{F}(p \cup q,(\mathcal{W}\sim p) \cup r) \end{array}\right\} \overset{c}{\Rightarrow} i \in \mathcal{F}(p \cup q,r) \quad (**)$$

$$\left.\begin{array}{l} \left.\begin{array}{l} \quad \boxed{i \in \mathcal{F}(p,q)} \\ \overset{c}{\Rightarrow}\ i \in \mathcal{F}(q,q) \end{array}\right\} \overset{v}{\Rightarrow} i \in \mathcal{F}(p \cup q,q) \\ \overset{**}{\Rightarrow} i \in \mathcal{F}(p \cup q,r) \end{array}\right\} \overset{iv}{\Rightarrow} i \in \mathcal{F}(q,r) \qquad \square$$

Recall that condition (iv) without the proviso is characterized by the principle of Strengthening the Antecedent. It will be clear that on Stalnaker's account this principle does not have general validity.
It may very well be that $\phi \rightarrow \psi$ is true, while $(\phi \wedge \chi) \rightarrow \psi$ is false. For, it may very well be that ψ is true in the ϕ-world most similar to the actual world, whereas ψ is false in the $(\phi \wedge \chi)$-world most similar to the actual world.
Formally:

II.52. EXAMPLE. Set $W = \{1,2\}$; $S_1(\emptyset) = S_2(\emptyset) = \emptyset$; $S_1(\{1\}) = S_1(\{1,2\}) = \{1\}$; $S_2(\{2\}) = S_2(\{1,2\}) = \{2\}$; $S_1(\{2\}) = \{2\}$; and $S_2(\{1\}) = \{1\}$.
Note that S is stalnakerlike. Let ϕ, ψ and χ be atomic sentences, and take $I(\phi) = \{1,2\}$; $I(\psi) = \{1\}$; and $I(\chi) = \{2\}$. Then $1 \in [\phi \rightarrow \psi]$, but $1 \notin [(\phi \wedge \chi) \rightarrow \psi]$. □

How much of the principle of Strengthening the Antecedent is retained in Stalnaker's theory? Or better, under what conditions *is* it allowed to infer $(\phi \wedge \chi) \rightarrow \psi$ from $\phi \rightarrow \psi$?

Note first that given our basic logic $\mathfrak{B}$ we will always be in one of the following situations:

(a) both $\phi \rightarrow \chi$ and $\phi \rightarrow \sim\chi$ are true
(b) $\phi \rightarrow \chi$ is true and $\phi \rightarrow \sim\chi$ is false
(c) $\phi \rightarrow \chi$ is false and $\phi \rightarrow \sim\chi$ is true
(d) both $\phi \rightarrow \chi$ and $\phi \rightarrow \sim\chi$ are false

$\mathfrak{B}$ does not allow us to infer $(\phi \wedge \chi) \rightarrow \psi$ from $\phi \rightarrow \psi$ in any of those situations. According to Stalnaker we may do so in each of these situations *except* in situation (c). His theory says that (d) will never in fact occur (apply CEM). And it also says that we may strengthen the antecedent with another consequent: in view of (iv) of proposition 51 the following argument form is valid

ASC $\quad \phi \rightarrow \psi, \phi \rightarrow \chi / (\phi \wedge \psi) \rightarrow \psi$.

Since $\phi \rightarrow \chi$ is true both in situation (a) and in situation (b), only situation (c) is left over.

The above suggests several descriptions of the logic determined by the class of stalnakerlike selection functions. From theorem 50, one might gather that it is just $\mathfrak{B}$ + CEM + CE + MP, i.e. the weakest logic containing all arguments of the forms CEM, CE, and MP. In view of proposition 51 we could restate this by saying that it is $\mathfrak{B}$ + AD + ASC + CEM + MP. But there is a surprise in store. In the next section it will appear that the logic determined by the class of stalnakerlike selection functions is not compact. So neither of these descriptions is enough.

II.3.2. Comparative similarity according to Lewis

II.53. DEFINITION. Let $\mathcal{D}$ be an accessibility function for $\mathcal{W}$. An ordering function for $\mathcal{W}$, given $\mathcal{D}$, is any function that assigns a strict partial order $<_i$ of $\mathcal{D}_i$ to each $i \in \mathcal{W}$. Whenever $j <_i k$ we say that j is *closer to* i *than* k. We write '$j \leqq_i k$' iff either $j <_i k$ or $j \in \mathcal{D}_i$ and $j = k$. □

The strict partial ordering $<_i$ is meant to play the role of a *comparative similarity* relation, excepting that as yet no constraints are imposed on it other than the ones one would impose on any relation that is to hold between three objects i, j and k iff j is more ... to i than k is. Below it will be our object to see if the relation $<_i$ will have to have other properties besides the irreflexivity and transitivity that make it a strict partial order.

II.54. DEFINITION. Let [ ] be a standard interpretation over a given set $\mathcal{W}$ of worlds. Let $\mathcal{D}$ be an accessibility function for $\mathcal{W}$, and < be an ordering function for $\mathcal{W}$ (given $\mathcal{D}$). [ ] is *based on* < iff for any two sentences ϕ and ψ,

$i \in [\phi \rightarrow \psi]$ iff for any $j \in [\phi] \cap \mathcal{D}_i$, there is some $k \in [\phi] \cap \mathcal{D}_i$ such that (i) $k \leqq_i j$; and (ii) $l \in [\psi]$ for any l such that $l \in [\phi] \cap \mathcal{D}_i$ and $l \leqq_i k$. □

The structure of this rather intricate truth condition is perhaps better displayed if we use predicate logical notation:

$$i \in [\phi \rightarrow \psi] \text{ iff } \forall j \in [\phi] \cap \mathcal{D}_i\ \exists k \in [\phi] \cap \mathcal{D}_i [k \leqq_i j\ \&\ \forall l \in [\phi] \cap \mathcal{D}_i (l \leqq_i k \supset l \in [\psi])]$$

So, for a sentence of the form 'if it had been the case that ϕ, then it would have been the case that ψ' to be true it is not strictly necessary that ψ holds at *all* accessible [ϕ]-worlds; if one can only find, for each accessible [ϕ]- world j some [ϕ]-world k closer to the actual world than j such that ψ holds at k and at all [ϕ]-worlds still closer to the actual world than k.

Part of the complexity of this definition is due to the fact that the ordering functions of definition 53 are not presupposed to satisfy the

Limit Assumption: for any $i \in \mathcal{W}$, $<_i$ is well-founded.

Call any $j \in \mathcal{W}$ a *closest* p-world to i iff $j \in p \cap \mathcal{D}_i$ and no $k \in p$ exists such that $k <_i j$. The Limit Assumption can then be restated as saying that for any proposition p such that $p \cap \mathcal{D}_i \neq \emptyset$ there is some closest p-world to i.[15)]

When the Limit Assumption is satisfied things turn out a lot less complicated.

II.55. PROPOSITION. Let $\mathcal{D}$ be an accessibility function for $\mathcal{W}$, and $<$ an ordering function. Suppose the Limit Assumption is satisfied. Then for any interpretation [ ] based on $<$ the following holds:

$i \in [\phi \rightarrow \psi]$ iff every closest $[\phi]$-world to i is an element of $[\psi]$.

PROOF. Left to the reader. □

We already saw why Lewis does not accept the Limit Assumption, and we will not find any good reason to disagree with him on this point. The same applies to this constraint:

Connectedness: for every $j,k \in \mathcal{D}_i$, either $j = k$, or $j <_i k$, or $k <_i j$.

To see that this is just a reformulation of Stalnaker's Uniqueness constraint in terms of ordering functions, consider $S_i(\{j,k\})$ for arbitrary i, j, k. Then if $S_i\{j,k\} \neq \emptyset$, i.e. if it is not absurd to entertain the counterfactual assumption that the world i might have been like j or like k, either $S_i(\{j,k\}) = \{j\}$ or $S_i(\{j,k\}) = \{k\}$, i.e. either j must be the $\{j,k\}$- world most similar to i, or k must be the $\{j,k\}$-world most similar to i. So if $j \neq k$, $j <_i k$ or $k <_i j$.

Lewis weakens Connectedness to

Almost-Connectedness: for any j, k, l $\in \mathcal{D}_i$, if j $<_i$ l, then either j $<_i$ k or k $<_i$ l.

To see the virtues of this constraint, let us write j $\simeq_i$ k iff neither j $<_i$ k not k $<_i$ j while both j and k belong to $\mathcal{D}_i$. Note that $\simeq_i$ is reflexive and symmetric, but not in general transitive. So we are not allowed to read j $\simeq_i$ k as j and k are *equally* similar to i.

II.56. PROPOSITION. Let < be an ordering function for $\mathcal{W}$, given the accessibility function $\mathcal{D}$. The following statements are equivalent.

(i) j $\simeq_i$ k iff for *any* l $\in \mathcal{D}_i$, if j $<_i$ l then k $<_i$ l and if l $<_i$ j then l $<_i$ k

(ii) $\simeq_i$ is transitive

(iii) $<_i$ is almost connected

So, how reasonable is it to assume that < is almost-connected? According to the above equivalences, just as reasonable as it is to assume that $\simeq_i$ is transitive. And how reasonable this is seems to depend on what you have in mind when you talk about comparative similarity.

Suppose, at one extreme, you like to think of $\simeq_i$ as meaning something like 'are indistinguishable to the naked eye'. Then cases where the transitivity of the relation breaks down can easily be found. You can for example imagine three (almost) identical triplet brothers, two of which are just distinguishable, while neither is distinguishable from the third. This could of course arise with any vague equivalence and there is a ready analogy with a notion such as 'looks equally long to the naked eye'.

Another extreme, you could for example think of $\simeq_i$ as something like the limit reached by a series of increasing precise ways of judging objects to be indistinguishable. This way you could perhaps hold onto the idea that $\simeq_i$ is transitive, just as 'equally long' could be kept transitive by taking increasingly precise measurements of lengths into consideration. The increasingly high resolution removes counterexamples to the transitivity by deciding that some things which earlier on looked equally similar to some other thing were not after all. By removing the vagueness, you hope to eliminate the intransitivity.

These are just two ways of thinking about comparative similarity, and are only intended to illustrate that differing opinions about whether $\simeq_i$ should be almost connected or not are defensible.

If we restate Stalnaker's Centering constraint in terms of ordering functions we get

Centering: $i \in \mathcal{D}_i$ for any $i \in \mathcal{W}$; and for any $j \in \mathcal{D}_i$, $i \leqq_i j$.

This condition is accepted by Lewis, believing as he does that there is only one world as similar to i as any world can be, namely i itself. One part of this - *Faithfulness* as we shall call it - seems very reasonable: no world j can be *closer* to i than i itself. But the other part - no world j can be *equally* close to i as i itself - seems less so. Again, it all depends on how you like to think of comparative similarity.

Presumably if you take all of the characteristics of the worlds into consideration in judging similarity relations, there will be just one world which resembles the real one as much as it itself does, that is to say, which is indistinguishable from it. But suppose you were to decide to judge similarity in terms of just some of these characteristics. Then there could obviously be other worlds which agree with the real one.

Finally we come to

Universality: for every $i \in \mathcal{W}$, $\mathcal{D}_i = \mathcal{W}$.

Suppose $\mathcal{D}$ is not universal. Note that according to definition 55, if $[\phi] \cap \mathcal{D}_i = \emptyset$, $i \in [\phi \rightarrow \psi]$ for *any* ψ. So if j happens to be inaccessible from i, this means so much as that from the point of view of i, it is absurd to suppose that the world might have been like j.

Is it reasonable to suppose that there are any such inaccessible possible worlds? Lewis (1981: 210) knows of no very strong reasons for or against it, and accordingly he treats Universality as an optional extra. We will not have any more to say on this than he.

II.57. PROPOSITION.

(i) Within the class of all ordering functions, all arguments of the form CI, CW, CC, AD, and ASC are valid.

(ii) An ordering function is almost-connected iff it validates all arguments of the form
ASP: $\phi \rightarrow \psi, \sim(\phi \rightarrow \sim\chi)/(\phi \wedge \chi) \rightarrow \psi$.

(iii) An ordering function is connected iff it validates all arguments of the form CEM.

(iv) An ordering function is faithful iff it validates all arguments of the form MP.

(v) An ordering function is centered iff it validates all arguments of the form MP as well as all arguments of the form CS.

PROOF. Straightforward. □

It is worth noticing that (i) of the above proposition remains true if we drop the requirement that $<_i$ be irreflexive for any i. In due course it will appear that there are no special argument forms corresponding to this property. If we drop the requirement that $<_i$ be transitive, it appears that only CI and CW remain valid; all three of CC, AD, ASC can be falsified on non-transitive 'ordering' functions, though not on just any non-transitive ordering function.[16)]

The principle ASP introduced in (ii) says that an antecedent ϕ of a counterfactual $\phi \rightarrow \psi$ may be strengthened with χ provided that the counterfactual assumption does not exclude the possibility that χ. So, given the validity of

ASC, there is only one case in which it is not allowed to strengthen the antecedent ϕ with χ, namely the case where $\phi \rightarrow \sim\chi$ is true and $\phi \rightarrow \chi$ is false. This is just what we found for Stalnaker's theory.

ASP is very attractive, intuitively speaking, and so far nobody every put forward a convincing counterexample.[17)] It is tempting to see this as evidence in favour of the almost-connectedness of the comparative similarity relation - or, if you believe that this relation just cannot be almost connected as evidence against any attempt of founding the interpretation of counterfactuals on a comparative similarity relation in the first place.

In the sequel we will refer to the logic $\mathfrak{B}$ + AD + ASC as $\mathfrak{P}$, and to $\mathfrak{P}$ + ASP as $\mathfrak{V}$. In section III.4 we will see that these logics are determined by the class of all ordering functions and the class of all almost connected ordering functions respectively.

In connection with (iv) and (v) it is worth noticing that all ordering functions validating CS also validate MP. Only if we drop the requirement that $<_i$ be irreflexive, it becomes possible to construct examples of ordering functions on which CS is valid but MP is not; because in that case it does not follow that for no $j \in \mathcal{D}_i$, $j <_i i$ if for any $j \in \mathcal{D}_i$, $i \leq_i j$.

II.58. DEFINITION. Let $\phi_1,\ldots,\phi_n,\ldots$ be countably many distinct atomic sentences, and let ψ_k and χ_k for any $k > 0$ be defined as

$$\psi_k = \sim((\phi_1 \vee \ldots \vee \phi_{k+1}) \rightarrow (\phi_1 \vee \ldots \vee \phi_k))$$
$$\chi_k = ((\phi_1 \vee \ldots \vee \phi_{k+1}) \rightarrow \sim(\phi_1 \vee \ldots \vee \phi_k))$$

Let Δ be the set of all ψ_k's and χ_k's.

LIM is to be the argument $\Delta/\bot$. □

II.59. PROPOSITION. Let < be an ordering function for $\mathcal{W}$, given $\mathcal{D}$. Then

< satisfies the Limit Assumption iff LIM is valid on <.

PROOF. Suppose first that < does not satisfy the Limit Assumption. Then we can find for i an infinite sequence of worlds $j_1,\ldots,j_k,\ldots$ such that $j_{k+1} <_i j_k$ for any k. Let I be any atomic interpretation such that $K(\phi_k) = \{j_k\}$ for any k. Consider the interpretation [ ] conforming to I and based on <. It is easy to see that $i \in [\psi_k]$ and $i \in [\chi_k]$ for any k. Hence LIM is not valid on <.

Conversely, suppose that LIM is not valid on <. Then we can find an interpretation [ ] based on < such that for some $i \in W$, $i \in [\psi_k]$ and $i \in [\chi_k]$ for any k. Note first that for every $k > 1$, $[\phi_1 \vee \ldots \vee \phi_{k+1}] \cap \mathcal{D}_i \neq \emptyset$. This is so because for every $k > 1$, $i \in [\psi_k]$. Furthermore, since $i \in [\chi_k]$ there must be for any $j \in [\phi_1 \vee \ldots \vee \phi_{k+1}] \cap \mathcal{D}_i$ a closest $[\phi_1 \vee \ldots \vee \phi_{k+1}]$-world j such that $j \notin [\phi_1 \vee \ldots \vee \phi_k]$. *A fortiori*, there will be for any $j \in [\phi_1 \vee \ldots \vee \phi_k]$ some $j' \in [\phi_1 \vee \ldots \vee \phi_{k+1}] \sim [\phi_1 \vee \ldots \vee \phi_k]$ such that $j' <_i j$. So for $j_1 \in [\phi_1]$ there is some $j_2 \in [\phi_1 \vee \phi_2] \sim [\phi_1]$ such that $j_2 <_i j_1$. For this world j_2 in its turn there is some $j_3 \in [\phi_1 \vee \phi_2 \vee \phi_3] \sim [\phi_1 \vee \phi_2]$ such that $j_3 <_i j_2$. Repeating this, we find an infinite sequence sequence $j_1,\ldots,j_k,\ldots$ such that for any k $j_{k+1} <_i j_k$, which means that < does not satisfy the Limit Assumption. □

II.60 COROLLARY. The logic determined by the class of all ordering functions that satisfy the Limit Assumption is not compact. Neither is the logic determined by the class of all (almost) connected ordering functions satisfying the Limit Assumption.

PROOF. Let Δ' be any finite subset of the set Δ introduced in definition 58. From the proof of proposition 59 it is obvious that we can find an ordering function satisfying the Limit Assumption on which Δ' is satisfiable. So $\Delta'/\bot$ does not belong to the logic determined by the class of all ordering functions satisfying the Limit Assumption. Also, it is clear that nothing changes if we restrict ourselves to the class of all almost-connected ordering functions or to the class of all connected ordering functions. □

Let $\mathcal{W}$ be a set of worlds, $\mathcal{D}$ an accessibility function for $\mathcal{W}$, $<$ an ordering function for $\mathcal{W}$, given $\mathcal{D}$, and [ ] an interpretation based on $<$. Call a sentence ϕ *entertainable* at i iff $[\phi] \cap \mathcal{D}_i \neq \emptyset$. Then, if $<$ satisfies the Limit Assumption, the following holds,

(*) If ϕ is entertainable at i, there is a world j such that $j \in [\psi]$ for any ψ such that $i \in [\phi \rightarrow \psi]$

That is, if ϕ is entertainable at i we can always find at least one world j at which *all* counterfactual consequences of ϕ hold. If $<$ does not satisfy the Limit Assumption we can no longer be certain of this. At best we have

(**) Let ϕ be entertainable at i, and let Γ be any *finite* subset of $\{\psi \mid i \in [\phi \rightarrow \psi]\}$. Then there is a world j such that $j \in [\psi]$ for any $\psi \in \Gamma$.

That is, one cannot always find a *possible* world j at which all the counterfactual consequences of ϕ hold. Sometimes they are *virtual* worlds - the ultrafilter extensions of $\{[\psi] \mid i \in [\phi \rightarrow \psi]\}$ - at which they all hold.

Several authors, in particular Pollock (1976, 1981) and Herzberger (1979), consider (*), on intuitive grounds, no less than an adequacy condition for any semantics of counterfactuals. Consequently they either think, as does Herzberger, that we must impose the Limit Assumption as a constraint on the comparative similarity relation, or they conclude, as does Pollock, that we must forget about using a comparative similarity relation as the basis for the interpretation of counterfactuals.

Lewis (1981: 229) thinks that despite the formal advantages of (*) it is best not to impose the Limit Assumption. I am inclined to agree with him, knowing that this is not the first time that an intuitive opinion on a general case (as formulated in (*)) turns out to hold only for the finite case (as formulated in (**)).

Still, the above leaves several important matters undecided. How exactly is the notion of comparative similarity to be understood? In what respects are the worlds to be compared? In all respects? In his (1973) Lewis steps pretty lightly over these questions. He simply advices us to take into account *over all* similaritics. He a¯nits that this is vague, but he says that 'it is vague - very vague - in a well-understood way ... it is just the sort of primitive that we must use to give a correct analysis of something that is itself undeniable vague' (see Lewis 1973: 91). In section 4 we will see how well this vagueness can be understood.

II.3.3. Premise semantics

Premise semantics, developed by Veltman (1976) and by Kratzer (1979, 1981), was like Stalnaker's theory inspired by Ramsey's advice for evaluating conditionals.
To recapitulate: first, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is then true.

Premise semantics starts off with an explication of this advice within possible-worlds semantics - an explication which is truer to the letter than that which Stalnaker gave. The following stepwise analysis will make this clear.

... your stock of beliefs ...

II.61. DEFINITION. A *premise function* for a given set W of possible worlds is any function P that assigns to each i in W a set P_i of propositions in W. The elements of P_i are called the *premises* pertaining to i.

For the time being you are asked to conceive of the elements of P_i as your beliefs about the world i. It could be that there are worlds i about which you do not believe anything, in which case $P_i = \emptyset$. Some of your beliefs about the world i may be incorrect, then there are propositions p such that $p \in P_i$ while $i \notin p$. And perhaps you believe incompatible things about some world i, in which case there will be no world j such that $j \in p$ for every $p \in P_i$. P is supposed to give a time slice through your beliefs; if you change them, P will have to change, too. Think of P_i as your present beliefs about i.

... add the antecedent ... whatever adjustments are required ...

Suppose you are evaluating the counterfactual 'if ϕ had been the case then ψ would have been the case' - should you accept or reject it? Suppose furthermore that your stock of beliefs P_i about the actual world i does not admit the proposition p expressed by ϕ: $\cap P_i \cap p = \emptyset$. Then, having added p to your stock of beliefs, you will have to make some adjustments in order to achieve consistency. P_i does not admit p, but perhaps some non-empty subset of P_i does. Now imagine a *maximal* non-empty subset of P_i which admits p, i.e. a set Q of propositions such that $\emptyset \neq Q \subset P_i$ and $\cap Q \cap p \neq \emptyset$. while for every proper extension Q' of Q within Q_i, $\cap Q' \cap p = \emptyset$. If you were to replace P_i by this set, and add p, then presumably you will have made the necessary changes.

Note that there may well be *more than one* maximal p-admitting subset of P_i, as is shown by the following example.

EXAMPLE. Let $\underset{\sim}{W} = \{1,2,3,4,5,6\}$; $P_1 = \{\{1,5,6\},\{1,3,5\},\{1,2,6\}\}$; $p = \{2,3,4\}$. Both $\{\{1,3,5\}\}$ and $\{\{1,2,6\}\}$ are maximal p-admitting subsets of P_1. (If you want to give some flesh and blood to this, see table I in section I.1.1.)

Worse, maybe there are *no* maximal p-admitting subsets of P_i at all, for either of two reasons.

(i) Perhaps no (non-empty) subset of P_i admits p; take for example $p = \emptyset$.

(ii) Or, even given that there are p-admitting subsets, it could turn out you could add premise after premise to these without ever reaching a limit. Suppose, for example, that you believe that there are at least n grains of salt, this for each natural number n. This stock of beliefs does not admit the proposition that there are finitely many grains. Nor does any of its infinite subsets. The only subsets which do so are the finite ones. As each finite subset will admit this proposition, however, none can be maximal.

Formally, this example looks like this:

EXAMPLE. Let $W = \omega + 1$. Take $p \in P_\omega$ iff $p = \{\sigma \mid \sigma \geqq \tau$ for some $\tau \in \omega\}$. It is easy to see that a non-empty subset of P_ω admits ω iff it is finite.

There is a very natural way of avoiding the difficulties involved in this phenomenon: we can resign ourselves to what we shall call the Finiteness Assumption and assume that you hold only a finite number of beliefs about each world. And there is another equally convenient but less natural way, too: just assume that there are no p-admitting subsets without a maximal extension:

II.62. DEFINTION. Let P be a premise function for W. P *satisfies the Finiteness Assumption* iff the cardinality of every set P_i is finite. P *satisfies the Limit Assumption* iff for any $p \subset W$ the following holds:

each p-admitting subset of every set P_i is a subset of some maximal p-admitting subset of P_i.

Obviously, any premise function satisfying the Finiteness Assumption will also satisfy the Limit Assumption.

... *consider whether or not the consequent is then true.*

II.63. DEFINITION.Let [ ] be a standard interpretation of a given language L over a given set of worlds W, and let P be a premise function for W. [ ] is *based on* P iff for every two sentences ϕ and ψ,

$i \in [\phi \rightarrow \psi]$ iff any non-empty $[\phi]$-admitting subset Q of P_i can be extended to a $[\phi]$-admitting subset Q' of P_i such that $\cap Q' \cap [\phi] \subset [\psi]$.

If P satisfies the Limit Assumption this reduces to

$i \in [\phi \rightarrow \psi]$ iff any non-empty maximal $[\phi]$-admitting subset Q of P_i is such that $\cap Q \cap [\phi] \subset [\psi]$.

It should be fairly clear that the definition follows Ramsey as far as he goes. But it goes further than he does by dealing with the eventuality that there are various ways of making whatever adjustments are required in order to maintain consistency. This is most easily seen if you make the Limit Assumption; then as we have seen there may be various different maximal $[\phi]$-admitting subsets. (Quantification over these has among its credentials that you do not get a conditional contradiction - both $\phi \rightarrow \psi$ and $\phi \rightarrow \sim\psi$ are true at i - except as a result of there being no non-empty $[\phi]$-admitting subsets of P_i, in which case $\phi \rightarrow \psi$ and $\phi \rightarrow \sim\psi$ are both vacuously true at i.)

Without the Limit Assumption, the formulation of the truth conditions becomes more complicated, but the idea remains the same: $\phi \rightarrow \psi$ is true at world i just in case one can add sufficiently many premises to any $[\phi]$-admitting subset of P_i to reach a point where one has got a $[\phi]$-admitting subset of P_i which in conjunction with $[\phi]$ implies $[\psi]$. That one may go on adding more and more premises, still keeping $[\phi]$-admitting subsets of P_i, is of minor importance.

If $\phi \rightarrow \psi$ is vacuously true, this may, in the extreme case, be due to ϕ's being absurd. (If $[\phi] = \emptyset$ there are no subsets of P_i admitting $[\phi]$.) Alternatively, it could result from your not being able to entertain ϕ without first rejecting everything else which you believed in. (Then the only $[\phi]$-admitting subset of P_i is empty.) The definition assumes, as it were, that you are not prepared to do this, that you find these sorts of ϕ's just as implausible as the absurdum. If you do not, then you could consider adapting the truth conditions to meet your taste simply by leaving out the requirement that the $[\phi]$-admitting subsets are not empty. Alternatively, you could adapt yourself to the truth conditions by choosing always to believe the trivial truth. This amounts to including W in each P_i. In fact you needn't even go this far, the simplest way to adapt your beliefs is to make sure that they are *universal* in the following sense.

Universality: P is *universal* iff for every i, $\bigcup P_i = W$.

You can verify this yourself by proving the following proposition.

II.64. PROPOSITION. Let P be a universal premise function for W. Suppose that [ ] is a standard interpretation of L over W based on P. Then $i \in [\phi \rightarrow \psi]$ iff each $[\phi]$-admitting subset Q of P_i can be extended to a $[\phi]$-admitting subset Q' of P_i such that no $\bigcap Q' \cap [\phi] \subset [\psi]$.

Another constraint you might want to impose on the premise functions P is this.

Faithfulness: $i \in p$ for every $p \in P_i$.

Indeed, it is arguable that the truth definition given above does not deserve the name if we do not impose this additional requirement. Without it, any phantasm, utterly unfounded, but nevertheless figuring on ones stock of beliefs, could make a difference to the truth values of ones counterfactual statements. But clearly that is absurd. Suppose you think that $\phi \rightarrow \psi$ is true, purely as a result of applying Ramsey's test to certain mistaken beliefs which you may have. And suppose you were at some later stage to discover that these beliefs are mistaken. Wouldn't you then want to revise your judgment of $\phi \rightarrow \psi$ and say it was false all along? By accepting Faithfulness we remove all erroneous beliefs from our premise sets. Now P_i contains only those propositions that we rightly believe to hold at i.

Many will think that imposing Faithfulness as a constraint on the premise functions P is still not enough to turn our definition into a decent truth definition. Erroneous beliefs may no longer interfere in matters of truth and falsity, but it may still happen that a counterfactual statement $\phi \rightarrow \psi$ owes its truth value simply to a lack of knowledge on the speaker's part. Suppose that, for all you know, $\phi \to \psi$ is 'false' at i, but that upon further investigation it appears that χ is true at i, and that adding $[\chi]$ to your stock of knowledge changes $\phi \to \psi$'s truth value. Wouldn't you then want to say that your earlier judgment was mistaken? It would seem that there is only one way to meet this criticism and that is to impose

Exhaustiveness: $P_i = \{p \mid i \in p\}$.

Unfortunately, imposing Exhaustiveness turns $\to$ into something very much like a strict implication.

II.65. PROPOSITION.[18)] Let P be an exhaustive premise function for W, and [ ] any standard interpretation based on P. Then for any i, ϕ, ψ the following holds.

(i) if $i \in [\phi]$ then $i \in [\phi \to \psi]$ just in case $i \in [\psi]$;

(ii) if $i \notin [\phi]$ then $i \in [\phi \to \psi]$ just in case $[\phi] \subset [\psi]$.

PROOF.

(i) Suppose $i \in [\phi]$. Then P_i extends every $[\phi]$-admitting subset O of P_i. Clearly $\cap P_i \cap [\phi] \subset [\psi]$ iff $i \in [\psi]$.

(ii) Suppose $i \notin [\phi]$. Let j be any $[\phi]$-world. (If there is no $[\phi]$-world there is nothing to prove.)Consider $Q = \{q \mid \{i,j\} \subset q\}$. Q is a maximal $[\phi]$-admitting subset of P_i. Clearly, $\cap Q \cap [\phi] \subset [\psi]$ iff $j \in [\psi]$. So $i \in [\phi \to \psi]$ iff for any $j \in [\phi]$ it holds that $j \in [\psi]$. □

This result is rather astonishing. We certainly have not gone through all these definitions just to end up with a truth definition that we already rejected in chapter II. So we'd better find some good reasons to reject Exhaustiveness.

II.3.4. Comparative similarity induced

Let P be a faithful premise function for W. Let Q and Q' be non-empty subsets of P_i. Suppose Q is a proper subset of Q'.

Now compare two worlds j and k, j an element of $\cap\mathcal{Q}'$, and k an element of $\cap\mathcal{Q}$ but not of $\cap\mathcal{Q}'$. Clearly j resembles i in some respects and k resembles i in some respects. Since $\mathcal{Q}$ is a proper subset of $\mathcal{Q}'$, we might even be inclined to say that j resembles i in more respects than k does. But then we realize that $\mathcal{Q}$ might have another extension $\mathcal{Q}'' \subset \mathcal{P}_i$ such that $k \in \cap\mathcal{Q}''$ and $j \notin \cap\mathcal{Q}''$. That is, there might be some other respects in which k is more like i than j. How should we balance off these respects against each other? One answer would be to say that we cannot balance them off. The worlds j and k are incomparable. In some respects j is closer to i than k, therefore k is not more similar than i. In some other respects k is closer to i than j, therefore j is not more similar to i than k. Also, it would be mistaken to say that j and k are equally similar to i. Only if j and k are like j in exactly the same respects could one say so. If you think this is the correct strategy, and if at the same time you maintain that $\mathcal{P}$ should be exhaustive, i.e. that we should compare all worlds in all respects, then any two worlds j and k will become incomparable as far as their similarity to any world i (different from j and k) is concerned. This is one of the results reported below.

II.66. DEFINITION. Let $\mathcal{P}$ be a premise function for $\mathcal{W}$, and $<$ an ordering function for $\mathcal{W}$, given $\mathcal{D}$.

$\mathcal{P}$ is *interchangeable* with $<$ iff the following two statements are equivalent for any i, p, q.

(i) any non-empty p-admitting subset $\mathcal{O}$ of $\mathcal{P}_i$ can be extended to a p-admitting subset $\mathcal{O}'$ of $\mathcal{P}_i$ such that $\cap\mathcal{O}' \cap p \subset q$;

(ii) $\forall j \in p \cap \mathcal{D}_i \; \exists k \in p \cap \mathcal{D}_i [k \leq_i j \;\&\; \forall l \in p \cap \mathcal{D}_i (l \leq_i k \supset l \in q)]$ □

It will be clear that interchangeable premise functions and ordering functions evaluate counterfactuals alike.

II.67. PROPOSITION.[19] Let $\mathcal{W}$ be a set of worlds. Every premise function for $\mathcal{W}$ is interchangeable with some ordering function and *vice versa*.

PROOF. I only state the relevant definitions leaving it to the reader to complete the proof.

Given $\mathcal{P}$, set $\mathcal{D}_i = \bigcup \mathcal{P}_i$ for any i; and define $<$ by $j <_i k$ iff

(i) for any $p \in \mathcal{P}_i$, if $k \in p$ then $j \in p$;

(ii) for some $p \in \mathcal{P}_i$, $j \in p$ and $k \notin p$.

Given $<$ and $\mathcal{D}$, we can derive $\mathcal{P}$ as follows:

$p \in \mathcal{P}_i$ iff $p = \{j \mid j \leq_i k\}$ for some $k \in \mathcal{D}_i$. □

Note in passing that each premise function is interchangeable with some unique ordering function. The converse is not true. (Suppose $\mathcal{P}$ and $\mathcal{P}'$ are the same except that $\mathcal{P}_i = \{\{i\},\{i,j,k\}\}$ while $\mathcal{P}_i = \{\{i,j\},\{i,k\}\}$. As Lewis (1981: 223) notes in either case the derived ordering function is the same.)

II.68. PROPOSITION. Let $\mathcal{P}$ be a premise function for $\mathcal{W}$. Suppose $\mathcal{P}$ is interchangeable with $<$.
If $\mathcal{P}$ is exhaustive, the $<_i = \emptyset$ while $\mathcal{D}_i = \mathcal{W}$ for every $i \in \mathcal{W}$.

PROOF. Omitted. □

It will be clear, then, that we cannot compare the worlds in all respects - if we do so, the comparative similarity relation becomes empty. We must neglect some respects, or we must at least balance them off. But which respects are unimportant enough to be neglected, and when is one respect more important than another? Compare the following well known examples

If this pen was made of copper, it would conduct electricity.

If this pen was made of copper, copper would not conduct electricity.

If you agree that the first of these sentences is true and the second false, you will presumably agree that worlds in which one or the other physical law is broken are less similar to the actual world than worlds in which some contingent facts do not obtain. In terms of premise functions: in adjusting your stock of premises you will not be prepared to give up a law just to keep a contingent fact.

Are we never prepared to give up a law? What about sentences starting with

If copper did not conduct electricity, then .....

Are these always vacuously true? No, I would say they are not. Only, in those contexts where they are not vacuously true (or for that matter false) the proposition expressed by 'copper conducts electricity' is not being treated as a law. Suppose you want to *test* whether copper conducts electricity. Then you will try to bring about a situation in which copper does not conduct. You have asked yourself 'what would be the case, if this law were not to hold?' and in looking for an answer to that question you may have taken some other laws for granted, but not the one you are willing to test. That is not to say, however, that in another context your attitude towards the very same proposition will be the same. There are many contexts where you will not be willing to give up the proposition in question. In those contexts you just do not want to reckon with the possibility that copper that copper might not conduct electricity. When you treat 'copper conducts electricity' as a law, it sets limits to your field of view.

I think that there are other kinds of propositions, besides the propositions which we ordinarily think of as natural laws, which can set limits to our field of view. Take the case of the marbles. Suppose you know exactly where the three marbles are; red in 1, blue in 2, and yellow in 2. Now compare the following two sentences:

If the red marble had been in box 2, one of the others would have been in box 1.

If the red marble had been in box 2, all three would have been in box 2.

All the time we have been treating the proposition that each box contains at least one marble as a law governing the distribution of the marbles. Therefore I think you will affirm the first and reject the second of these sentences.

The role which laws - and other propositions we treat as such - play is important, since they determine which possible worlds can enter into the relation of comparative similarity and which cannot. Only those worlds in which the same laws hold as in the actual one can. But the role which laws play is not decisive. They do not determine which of these worlds are *more* similar to the actual world than which others.

There must be other factors which do, other characteristics with respect to which worlds can agree or differ. As we saw, not all characteristics of the worlds can be taken into account. I used to think (see Veltman (1976)) that just those characteristics should be taken into account with which one is acquainted, and I thought of each P_i as containing just these. I still think that those characteristics with which we are acquainted are the only ones which can matter, but I no longer think that they are all relevant. Paul Tichy (1976: 271) made this clear with the following example:

"Consider a man - call him Jones - who is possessed of the following dispositions as regards wearing his hat. Bad weather invariably induces him to wear his hat. Fine weather, on the other hand, affects him neither way: on fine days he puts his hat on or leaves it on the peg, completely at random. Suppose, moreover, that actually the weather is bad, so Jones *is* wearing his hat....".

Tichy then asks us to evaluate the sentence
If the weather were fine, Jones would be wearing his hat.
We know that Jones actually is wearing his hat. We also know that it is raining. Now we must add the proposition that the weather is fine to our stock of premises, thereby making whatever adjustments *minimally* required to retain consistency. Clearly, this can be done without Jones having to take his hat off. So, applying our premise semantical recipe, we see that the sentence in question is true. But obviously it is not.

Tichy's criticism was directed against Lewis' and Stalnaker's theories, but it applies to premise semantics just as well. I think that his example definitely shows that even some of the *known* characteristics of the actual world are irrelevant in assessing which worlds resemble it more closely than which others, but I would not know how to go about identifying the characteristics which *are* relevant. I am afraid that I must admit - together with Stalnaker, Kratzer, and to some lesser extent Lewis[20] - that everything depends on the context.

II.4. MODELTHEORETIC RESULTS

What is the logic determined by the class of faithful premise functions? Or, what is the logic determined by the class of all almost-connected and centered ordering functions? And what is the logic of the class of all universal stalnakerlike selection functions? In the previous sections we have encountered many different classes, all raising questions like the above. In this chapter we shall answer as many of them as we can. Fortunately, we will not have to treat them all separately. For example, we already know that premise functions and ordering functions are interchangeable. Furthermore, all the special properties these functions can have (think of faithfulness, universality and even the limit assumption) carry over from one kind to the other. So we can restrict ourselves to investigating just one of these classes. Since ordering functions are much easier to get to grips with than premise functions the choice is easily made.

II.4.1. Comparing ordering functions and selection functions

In section 2, the strategy followed in answering the above questions was this: let C be the class of ...-functions concerned. Firstly, determine the class K of all connectors interchangeable with some element of C; secondly, delineate the smallest *characterizable* class K' of connectors such that $K \subset K'$; finally, show that the weakest logic containing the argument forms characterizing K' is complete. Here I will often follow a different strategy, sometimes because there is a shortcut, sometimes because I can only find a roundabout way. For example, I have been unable to find a useful characterization of the class K of all connectors that are interchangeable with some ordering function - none, that is, that can serve as a starting point for a *direct* proof that the class of all connectors validating the logic $\mathbf{P}$ is the smallest characterizable class K' such that $K \subset K'$. What I can give, however, is a useful characterization of the class of all selection functions that are interchangeable with some ordering function; and this will at least make an indirect proof of the mentioned result possible.

The following proposition explains why this restriction to selection functions makes things a lot clearer.

II.70. PROPOSITION. Let F be a connector for W, and $<$ an ordering function for W. Suppose F is interchangeable with $<$. Then

(i) F is conclusive;

(ii) F is strongly conclusive iff $<$ satisfies the Limit Assumption

PROOF. (i) comes as no surprise.

(ii) Suppose $<$ satisfies the Limit Assumption. We must show that $\cap C_i(p) \in C_i(p)$ for C the consequence function derived from F. Let $M_i(p)$ for every $i \in W$ and $p \subset W$ be the set of p-closest worlds to i. Note that $M_i(p) \in C_i(p)$, whence $\cap C_i(p) \subset M_i(p)$. Given the interchangeability of F and $<$ we also have that $M_i(p) \subset q$ for every $q \in C_i(p)$. Hence $M_i(p) \subset \cap C_i(p)$. It is clear, then, that $\cap C_i(p) \in C_i(p)$.

To prove the converse, suppose that $\cap C_i(p) \in C_i(p)$ for every $i \in W$ and $p \subset W$. Consider any p such that $p \cap \mathcal{D}_i \neq \emptyset$. We must show that there is some p-closest world. Note first that $\cap C_i(p) \neq \emptyset$. Otherwise, the interchangeability of F and $<$ would yield

$$\forall j \in p \cap \mathcal{D}_i \; \exists k \in p \cap \mathcal{D}_i (k \leqq j \;\&\; \forall l \in p \cap \mathcal{D}_i (l \leqq k \supset l \in \emptyset))$$

which would mean that $p \cap \mathcal{D}_i = \emptyset$. Now, take any $m \in \cap C_i(p)$; we will show that m is a p-closest world to i. Assume for contradiction that there is some $n \in p \cap \mathcal{D}_i$ such that $n <_i m$. Then

$$\forall j \in p \cap \mathcal{D}_i \; \exists k \in p \cap \mathcal{D}_i (k \leqq j \;\&\; \forall l \in p \cap \mathcal{D}_i (l \leqq k \supset l \in W \sim \{m\})).$$

By the interchangeability of F and $<$ it follows that $W \sim \{m\} \in C_i(p)$. In other words $m \notin \cap C_i(p)$. Contradiction. □

Strongly conclusive connectors - or selection functions for that matter - are much easier to handle mathematically than 'weakly' conclusive connectors. And so are ordering functions which satisfy the Limit Assumption, as compared with ordering functions which do not.

Still, things will turn out complicated enough. This already becomes apparent if we just state - the proof is postponed - the theorem that tells which selection functions are interchangeable with some ordering function.

II.71. THEOREM. Let S be a selection function for W. S is interchangeable with some ordering function for W iff

(i) S validates AD, i.e. for any $p,q,r \subset W$, if $S_i(q) \subset p$ and $S_i(r) \subset p$ then $S_i(q \cup r) \subset p$

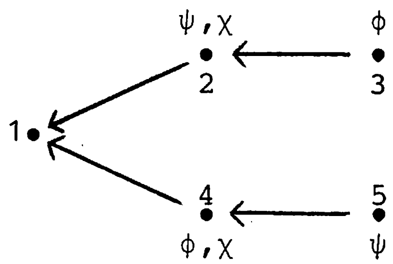
(ii) S validates ASC, i.e. for any $p,q,r \subset W$ if $S_i(q) \subset p$ then $S_i(q \cap r) \subset p$ *provided that* $S_i(q) \subset r$

(iii) S is *coherent*, i.e. if $Q \neq \emptyset$ and $j \in S_i(q)$ for every $q \in Q$, then $j \in S_i(\cup Q)$

This coherence condition is monstrous. Before we take a closer look, it is worth noticing that the following argument scheme is not generally valid within the class of ordering functions:

$$(*) \quad (\phi \vee \psi) \rightarrow \chi \;/\; (\phi \rightarrow \chi) \vee (\psi \rightarrow \chi)$$

EXAMPLE. Let $W = \{1,2,3,4,5\}$. Let $<$ be any ordering function for W such that $D_1 = W$ and $<_1 = \{\langle 1,2\rangle,\langle 1,3\rangle,\langle 1,4\rangle,\langle 1,5\rangle,\langle 2,3\rangle,\langle 4,5\rangle\}$. Let ϕ, ψ, χ be atomic sentences and set $I(\phi) = \{3,4\}$, $I(\psi) = \{2,5\}$ and $I(\chi) = \{2,4\}$. Let $[\ ]$ be the interpretation conforming to I and based on $<$. From the picture it is easily seen that $1 \in [(\phi \vee \psi) \rightarrow \chi]$, while neither $1 \in [\phi \rightarrow \chi]$, nor $1 \in [\psi \rightarrow \chi]$.



□

The example is typical for strict partial orderings. If $<$ is almost connected, things are different.

II.72. PROPOSITION. (*) is valid within the class of almost connected ordering functions.

PROOF. Omitted. □

Let us now turn to the coherence condition. Note that if the domain W of S is finite, S is coherent just in case S has the following property:

(iv) if $S_i(q \cup r) \subset W \sim \{j\}$ then $S_i(q) \subset W \sim \{j\}$ or $S_i(r) \subset W \sim \{j\}$

What we have here is a special case of

(v) if $S_i(q \cup r) \subset p$ then $S_i(q) \subset p$ or $S_i(r) \subset p$.

The argument scheme corresponding to (v) is (*). As we have seen, (*) does not hold on arbitrary ordering functions. So (v) will not generally hold for selection functions interchangeable with ordering functions. Apparently, the weaker condition (iv) does hold for these selection functions. But as we will see, (iv) is not characterizable. (Notice that (iv) would be characterizable if in our formal languages we could say something like: 'if q had been the case, the world would differ from j'.)

There are selection functions satisfying the conditions (i), (ii) and (iv) which lack the property laid down in (iii). Of course, these selection functions all have an infinite domain.

EXAMPLE. Let $\mathcal{W} = \omega$. Let S be the selection function for $\mathcal{W}$ given by the following clauses

(a) if p is infinite then $S_i(p) = \{n \in p \mid n \geq i\}$

(b) if p is finite then $S_i(p) = \{n \in p \mid \sim\exists m \in p(n < m < i)\}$

It is left to the reader to check that S satisfies conditions (i), (ii) and (v) (!!). It is clear then, that the following holds for every (non-empty) *finite* set $\mathcal{Q}$ of propositions in $\mathcal{W}$,

if $j \in S_i(q)$ for every $q \in \mathcal{Q}$ then $j \in S_i(\cup\mathcal{Q})$

Now consider the infinite set $\mathcal{Q}'$ of propositions given by

$q \in \mathcal{Q}'$ iff $q = \{k \in \omega \mid k \leq n\}$ for some $n \in \omega$.

Then $1 \in S_2(q)$ for every $q \in \mathcal{Q}'$, but $1 \notin S_2(\cup\mathcal{Q})$. □

Finally, we do have

II.73. PROPOSITION. Every selection function validating AD, ASC and ASP is coherent.

PROOF. Suppose $\mathcal{Q} \neq \emptyset$ and $j \in S_i(q)$ for every $q \in \mathcal{Q}$. Note first that $S_i(\cup\mathcal{Q}) \neq \emptyset$. Otherwise, we would have that for every $q \in \mathcal{Q}$,

$$\left.\begin{array}{l} S_i(\cup\mathcal{Q}) \subset q \\ S_i(\cup\mathcal{Q}) \subset \mathcal{W} \sim q \end{array}\right\} \overset{ASC}{\Rightarrow} S_i(\cup\mathcal{Q} \cap q) = S_i(q) \subset \mathcal{W} \sim q \Rightarrow S_i(q) = \emptyset.$$

Contradiction.

Since $S_i(\cup\mathcal{Q}) \neq \emptyset$, there will be some $q \in \mathcal{Q}$ such that $S_i(\cup\mathcal{Q}) \cap q \neq \emptyset$. Now assume $j \notin S_i(\cup\mathcal{Q})$. We then have

$$\left.\begin{array}{l} S_i(\cup\mathcal{Q}) \subset W \sim \{j\} \\ S_i(\cup\mathcal{Q}) \cap q \neq \emptyset \end{array}\right\} \overset{\text{ASP}}{\Rightarrow} S_i(\cup\mathcal{Q} \cap q) = S_i(q) \subset W \sim \{j\}.$$

Contradiction. □

These considerations help us to understand why it is often very cumbersome to prove modeltheoretic results for $\mathfrak{P}$ whereas the analogous proof for $\mathfrak{D}$ runs smoothly. Selection functions validating $\mathfrak{D}$ are coherent all by themselves, whereas selection functions validating $\mathfrak{P}$ have to be made coherent.

Here is the proof of the easy half of theorem 71. Suppose S is interchangeable with $<$. We must show that S satisfies conditions (i), (ii) and (iii). (i) and (ii) are taken for granted. To prove (iii), consider any non-empty set of propositions $\mathcal{Q}$ such that $j \in S_i(q)$ for every $q \in \mathcal{Q}$. We know from proposition 70 that $<$ satisfies the Limit Assumption. Let $M_i(q)$ for every $q \in \mathcal{Q}$ be the set of q-closest worlds to i. From the interchangeability of S and $<$ it follows that $j \in M_i(q)$ for every $q \in \mathcal{Q}$. So, if $k < j$, there will be *no* $q \in \mathcal{Q}$ such that $k \in q$. This means that $j \in M_i(\cup\mathcal{Q})$, i.e. the set of $\cup\mathcal{Q}$-closest worlds to i. By the interchangeability of $<$ and S, it follows that $j \in S_i(\cup\mathcal{Q})$.

The proof of the difficult half of theorem 71 runs as folllows.
Suppose S satisfies conditions (i), (ii), and (iii).
Let $\mathcal{D}_i$ for every $i \in W$ be the set of all j such that $j \in S_i(p)$ for some p.
Take $s_i(j)$ for every $j \in \mathcal{D}_i$ to be $\cup\{p \mid j \in S_i(p)\}$.
Since S is coherent $j \in S_i(s_i(j))$.
Now consider the ordering function $<$ for W given by

$k <_i j$ iff $k, j \in \mathcal{D}_i$, $s_i(j) \subset s_i(k)$, and $k \notin s_i(j)$.

Clearly, $<_i$ is transitive and irreflexive.
So it remains to be proved that S and $<$ are interchangeable.

Omitting the subscript i, we note first

(a) if $j \in S(p)$ then there is no $k < j$ such that $k \in p$. For, if $j \in S(p)$, then $p \subset s(j)$, and if $k < j$ then $k \notin s(j)$.

(b) if $j \in p \cap \mathcal{D}$ and $j \notin S(p)$, then there is some $k < j$ such that $k \in S(p)$.

Proof: first we show that $S(s(j) \cup p) \sim s(j) \neq \emptyset$. Assume for contradiction that $S(s(j) \cup p) \subset s(j)$. We then have

$$\left.\begin{array}{l} S(p) \subset (\mathcal{W} \sim p) \cup S(p) \\ S((\mathcal{W} \sim p) \cap s(j)) \subset (\mathcal{W} \sim p) \cup S(p) \end{array}\right\} \overset{AD}{\Rightarrow} S(s(j) \cup p) \subset (\mathcal{W} \sim p) \cup S(p) \quad (*)$$

$$\left.\begin{array}{l} S(s(j) \cup p) \subset (\mathcal{W} \sim p) \cup S(p) \quad (*) \\ S(s(j) \cup p) \subset s(j) \end{array}\right\} \overset{ASC}{\Rightarrow} S(s(j)) \subset (\mathcal{W} \sim p) \cup S(p)$$

This contradicts the fact that $j \in S(s(j))$, $j \in p$, $j \notin S(p)$.

Next we show that $S(s(j) \cup p) \sim s(j) \subset S(p)$.

$$\left.\begin{array}{l} S(p) \subset S(p) \cup s(j) \\ S(s(j)) \subset S(p) \cup s(j) \end{array}\right\} \overset{AD}{\Rightarrow} S(s(j) \cup p) \subset S(p) \cup s(j)$$

So, consider any $k \in S(s(j) \cup p) \sim s(j)$. Clearly $s(j) \subset s(k)$, $k \notin s(j)$. Hence $k < j$.

This completes the proof of (b).

(a) and (b) not only say that $<$ satisfies the Limit Assumption, but also that the set of p-closest worlds always coincides with $S(p)$. It is obvious, then, that S and $<$ are interchangeable.[21)] □

There are many variants to theorem 71. If you want to know which selection functions are interchangeable with some faithful/centered/universal ordering function then all you need to do is to add the condition that S itself be faithful/centered/universal. It is easy to show that $<$ is so defined that a selection function with any of these properties is transformed into an ordering function with the corresponding properties. Also the cases of almost-connected and connectedness can be handled in this manner. Only the proofs are somewhat more involved.

II.74. THEOREM. A selection function is interchangeable with some almost-connected/connected ordering function iff S validates AD, ASC, and ASP/CEM.

PROOF. Given proposition 73, we can be certain that S is coherent. So, let $<$ be defined as in the proof of theorem 71.
Suppose S validates ASP. We must show that $<$ is almost connected. Assume that $k < j$ and let l be any element of $\mathcal{D}$. (Like before, we omit the subscript i.)
Two possibilities obtain.
(a) $S(s(k) \cup s(l)) \cap s(l) = \emptyset$
In this case we have

$$\left.\begin{array}{l} S(s(k) \cup s(l)) \subset s(k) \\ S(s(k) \cup s(l)) \subset S(s(k) \cup s(l)) \end{array}\right\} \overset{\text{ASC}}{\Rightarrow} S(s(k)) \subset S(s(k) \cup s(l))$$

It follows that $k \in S(s(k) \cup s(l))$, whence $s(l) \subset s(k)$.
It also follows that $k \notin s(l)$. So, $k < l$.
(b) $S(s(k) \cup s(l)) \cap s(l) \neq \emptyset$.
Then we have

$$\left.\begin{array}{l} S(s(k) \cup s(l)) \cap s(l) \neq \emptyset \\ S(s(k) \cup s(l)) \subset S(s(k) \cup s(l)) \end{array}\right\} \overset{\text{ASP}}{\Rightarrow} S(s(l)) \subset S(s(k) \cup s(l))$$

This means that $l \in S(s(k) \cup s(l))$, whence $s(k) \subset s(l)$.
Since $s(j) \subset s(k)$, we see that $s(j) \subset s(l)$. So if we show that $l \notin s(j)$ we have proved that $l < j$.
Note first that $S(s(l) \cup s(j)) \cap s(j) = \emptyset$. Otherwise we could prove - using ASP, just like above - that $s(l) \subset s(j)$; but then it would follow that $s(k) \subset s(j)$, which contradicts the fact that $k \notin s(j)$. Given that $S(s(l) \cup s(j)) \cap s(j) = \emptyset$, it follows that $S(s(l) \cup s(j)) \subset s(l)$. Applying ASC we find $S(s(l)) \cap s(j) = \emptyset$. So $l \notin s(j)$.

Now suppose that S validates CEM. We must show that $<$ is connected. So, let k, j be any two elements of $\mathcal{D}$ such that $k \neq j$. This time we distinguish three cases.
(a)' $S(s(k) \cup s(j)) \cap s(j) = \emptyset$. Proceeding like under (a) we find that $k < j$.

(b)' $S(s(j) \cup (s(k)) \cap s(k) = \emptyset$. Similarly, we find that $j < k$.

(c)' $S(s(k) \cup s(j)) \cap s(j) \neq \emptyset$, and $S(s(k) \cup s(j)) \cap s(k) \neq \emptyset$. Proceeding like under (b), it can be shown that $s(k) \subset s(j)$, and $s(j) \subset s(k)$. From this it follows that $S(s(k)) = S(s(j))$. Using CEM, it is easy to see that $S(s(k)) = \{k\}$, and $S(s(j)) = \{j\}$. Hence $k = j$. □

II.4.2. The completeness of $\mathfrak{P}$

With the next two definitions we take the first steps on our way to a proof that the property of coherence is not characterizable.

II.75. DEFINITION. Let S be a selection function for W. S is *separative* iff for any $i,j \in W$ and $p,q \subset W$,

$$\text{if } j \in S_i(p) \text{ and } j \in S_i(q)\text{, then } p = q$$ □

Trivially, separative selection functions are coherent.

II.76. DEFINITION. Let S be a selection function for W. Consider the set U of all pairs $\langle i,p \rangle$ with $i \in W$, $p \subset W$, and $i \in p$. Consider also the function $^+$ from pow W into pow U defined by

$$p^+ = \{u \in U \mid u_1 \in p\}$$

Let T be any selection function for U. T is a *propositional dissection* of S iff for every $p \subset W$ and $u,v \in U$ the following holds

$$v \in T_u(p^+) \text{ iff } v_1 \in S_{u_1}(p) \text{ and } v_2 = p$$ □

Proofs to the effect that a given property is not characterizable often take the form of a recipe. They tell how a structure without the property concerned can be transformed into a structure with the property concerned in such a way that any set of sentences satisfiable on the original structure is also satisfiable on the derived one.

This case is no exception to that rule as the following propositions show.

II.77. LEMMA. Every selection function has a separative propositional dissection.

PROOF. Let S be a selection function for W. Let U and $^+$ be defined as above. Consider the selection function T for U given by the following clauses.

$$T_u(P) = \begin{cases} \{v \in U \mid v_1 \in S_{u_1}(p) \text{ and } v_2 = p\}, \text{ if } P = p^+ \\ \emptyset, \text{ otherwise} \end{cases}$$

Clearly T is separative and a propositional dissection of S. □

II.78. LEMMA. Let S be a selection function for W, and T a propositional dissection of S. Every set of sentences satisfiable on S is satisfiable on T.

PROOF. Take U and $^+$ as above. Let I be any atomic interpretation over W, and consider the atomic interpretation J over U that is related to S as follows:

$$J(\phi) = S(\phi)^+$$

We show, by induction on the complexity of ϕ, that the interpretation [ ] conforming to I and based on S, and the interpretation ⟦ ⟧ conforming to J and based on T are likewise related, i.e.

$$\text{for every sentence } \phi,\ [\![\phi]\!] = [\phi]^+$$

The case that $\phi = \sim\psi$ as well as the case that $\phi = (\psi \wedge \chi)$ can straightforwardly be proved once it has been checked that $^+$ is a boolean homomorphism from pow W into pow U. In particular we have

$$(W \sim p)^+ = U \sim p^+$$

and

$$(p \cap q)^+ = p^+ \cap q^+$$

So it remains to prove the case that $\phi = (\psi \rightarrow \chi)$. Suppose that $u \in [\![\psi \rightarrow \chi]\!]$. This is so iff $T_u([\![\psi]\!]) \dot{\subset} [\![\chi]\!]$.

By the induction hypothesis this is equivalent to:

$T_u([\psi]^+) \subset [\chi]^+$. Since T is a propositional dissection of S this can be equated with

$v_1 \in [\chi]$ for every v such that $v_1 \in S_{u_1}([\psi])$ and $v_2 = [\psi]$.
In other words, $v_1 \in [\chi]$ for every $v_1 \in S_{u_1}([\psi])$.
That is, $u_1 \in [\psi \to \chi]$.
Or, $u \in [\psi \to \chi]^+$. □

II.79. COROLLARY. The logic determined by the class of separative selection functions, and the logic determined by the class of coherent selection functions both equal $\mathbf{B}$.

Caution! From this corollary it does not follow that the logic determined by the class of all coherent selection functions validating AD and ASC is just $\mathbf{B}$ + AD + ASC = $\mathbf{P}$. From theorem 71 we know that the logic determined by this class equals the logic determined by the class of those ordering functions that satisfy the Limit Assumption. And as we saw when we discussed the Limit Assumption, the latter logic is not compact.

Still, lemmas 77 en 78 will prove to be of great help in the proof that the logic determined by the class of all ordering functions (those not satisfying the Limit Assumption included) *is* $\mathbf{P}$. And so will the next proposition.

II.80. THEOREM. Let S be a separative selection function for W. Suppose [ ] is a standard interpretation over W based on S which *verifies* $\mathbf{P}$. Then [ ] can be based on an ordering function for W.

PROOF. Let D_i for every $i \in W$ be the set of all j such that $j \in S([\phi])$ for some ϕ. Take $p_i(j)$ for every $j \in D_i$ to be the unique proposition p such that $j \in S_i(p)$. Since S is separative this is well defined. Now consider the ordering function < for W given by

$k <_i j$ iff $k, j \in D_i$, $p_i(j) \subset p_i(k)$, and $k \notin p_i(j)$.

Clearly $<_i$ is transitive and irreflexive for every i. (So far the proof is almost identical to the proof of theorem 71.) Keep in mind, however, that S does not necessarily *validate* AD and ASC. Therefore, we cannot expect < to satisfy the Limit Assumption. Another consequence is that now all obstacles have to be taken syntactically, where in the proof of theorem 71, we could freely base our arguments on the structural impact of AD and ASC.)

It remains to be proved that [ ] can be based on <. To this end it suffices to show that $i \in [\phi \rightarrow \psi]$ iff

$\forall j \in \mathcal{D}_i \cap [\phi]\ \exists k \in \mathcal{D}_i \cap [\phi]\ (k \leqq_i j\ \&\ \forall l \in \mathcal{D}_i \cap [\phi]\ \ (l \leqq_i k \supset l \in [\psi]))$.

Below the subscript i will again be omitted.

In one direction the proof is easy.

Suppose $S_i([\phi]) \not\subset [\psi]$. Consider any $j \in S([\phi]) \sim [\psi]$. Since [ ] verifies $\mathcal{P}$, $j \in [\phi]$. Obviously, $j \in \mathcal{D}$. If $k < j$, then $k \notin [\phi]$. For if $k < j$, then $k \notin p(j)$ and $p(j) = [\phi]$. In sum,

$\exists j \in \mathcal{D} \cap [\phi]\ \forall k \in \mathcal{D} \cap [\phi]\ (k \leqq j \supset \exists l \in \mathcal{D} \cap [\phi]\ (l \leqq k\ \&\ l \notin [\psi]))$.

For the converse, suppose that $S([\phi]) \subset [\psi]$.

Consider any $j \in \mathcal{D} \cap [\phi]$. We must show

$\exists k \in \mathcal{D} \cap [\phi]\ (k \leqq j\ \&\ \forall l \in \mathcal{D} \cap [\phi]\ (l \leqq k \supset l \in [\psi]))$.

There are two possibilities.

(a) $\forall l \in \mathcal{D} \cap [\phi]\ (l \leqq j \supset l \in [\psi])$. In this case there is nothing to prove.

(b) $\exists l_o \in \mathcal{D} \cap [\phi]\ (l_o \leqq j\ \&\ l_o \notin [\psi])$. Since $l_o \in \mathcal{D}$, $l_o \in S([\chi])$ for some χ.

Claim: there is some k_o such that $k_o \in S([\phi \vee \chi])$ and $k_o \notin [\chi]$.

Proof of the claim: as the following quasi derivation in which all applications of Replacement are left out shows, the argument form

$$\phi \rightarrow \psi,\ (\phi \vee \chi) \rightarrow \chi\ /\ \chi \rightarrow (\sim\phi \vee \psi)$$

belongs to $\mathcal{P}$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\phi}{(\sim\phi \wedge \chi) \to (\sim\phi \wedge \chi)}\,\text{CI}}{(\sim\phi \wedge \chi) \to (\sim\phi \vee \psi)}\,\text{CW} \qquad \dfrac{\phi \to \psi}{\phi \to (\sim\phi \vee \psi)}\,\text{CW}}{(\phi \vee \chi) \to (\sim\phi \vee \psi)}\,\text{AD} \qquad (\phi \vee \chi) \to \chi}{\chi \to (\sim\phi \vee \psi)}\,\text{ASC}$$

Now assume for contradiction that there is no k such that $k \in S([\phi \vee \chi]) \sim [\chi]$. Then we have $i \in [\phi \to \psi]$, $i \in [(\phi \vee \chi) \to \chi]$, and therefore also $i \in [\chi \to (\sim\phi \vee \psi)]$. This means that $S([\chi]) \subset ((\mathcal{W} \sim [\phi]) \cup [\psi])$. However, we already know that $l_o \in S([\chi])$, $l_o \in [\phi]$ and $l_o \notin [\psi]$. Contradiction.
Since $p(l_o) = [\chi] \subset [\phi \vee \chi] = p(k_o)$ and $k_o \notin [\chi]$, we find that $k_o < l_o$. Furthermore, since $\phi \to \psi \,/\, \phi \vee \chi \to \psi \vee \chi$ belongs to $\mathcal{P}$, we have that $S([\phi \vee \chi]) \sim [\chi] \subset [\psi]$. So, $k_o \in [\psi]$. Finally, for all $l < k_o$, it holds that $l \notin [\phi]$. So we see that $\forall l \in \mathcal{D} \cap [\phi]\ (l \leq k_o \supset l \in [\psi])$, which means that we are through. □

When I first mentioned separative selection functions I did not give any special reason for introducing them.
I think I can now fill up this gap by pointing out that the above proof really hinges on the assumption that S is separative rather than coherent. In this connection it is instructive to see where the proof goes wrong if you try it on a coherent selection function S.
Take $p_j = \cup\{[\phi] \mid j \in S([\phi])\}$ instead of the unique $[\phi]$ such that $j \in S([\phi])$. Everything works out well until you arrive at case (b). The problem is that for *infinite* $\mathcal{W}$, you can no longer be certain that for every $l \in \mathcal{D}$, $p(l) = [\chi]$ for some χ. And that is something really needed to carry through the rest of the proof.

We are ready now to prove the main result of this section.

II.81. THEOREM. $\mathfrak{P}$ is complete.

PROOF. Since $\mathfrak{P}$ is compact it is sufficient to show that $\mathfrak{P}$ is canonical. Let [ ] be any standard compositional interpretation verifying $\mathfrak{P}$. Let W be the set of worlds over which [ ] is defined. We must show that [ ] can be based on some $F \in K_{\mathfrak{P}}$.

Since $\mathfrak{B}$ is canonical, and $\mathfrak{B} \subset \mathfrak{P}$ we may without loss of generality assume that [ ] can be based on some conclusive connector $G \in K_{\mathfrak{B}}$.

Let U be the set of ultrafilters on W, and consider the function * from pow W into pow U given by $p^* = \{u \in U \mid p \in u\}$. By theorem 33, there is a strongly conclusive ultrafilter extension G' of G with U as its domain. Let $[\![\ ]\!]$ be the interpretation over U given by $[\![\phi]\!] = [\phi]^*$. By lemma 31, $[\![\ ]\!]$ is a standard interpretation that can be based on G'.

According to theorem 29 G' is interchangeable with some selection function S. $[\![\ ]\!]$ can be based on S.

Next consider $V = \{\langle u,P\rangle \mid u \in P \text{ and } P \subset U\}$. Let $^+$ be the function from pow U to pow V given by $P^+ = \{v \in V \mid v_1 \in P\}$. By lemma 77, we can find a separative propositional dissection T of S with V as its domain. By lemma 78, the interpretation $[\![\![\ ]\!]\!]$ given by $[\![\![\phi]\!]\!] = [\![\phi]\!]^+$ is a standard interpretation that can be based on T.

Now we can apply theorem 80: there is an ordering function < for V on which $[\![\![\ ]\!]\!]$ can be based. This function, in its turn, is interchangeable with a connector G'' for V. $G'' \in K_{\mathfrak{P}}$.

From G'' we can derive the desired connector F for W as follows. Set

$$i \in F(p,q) \text{ iff } \langle u(i),U\rangle \in G''(p^{*+},q^{*+})$$

(Here u(i) is the ultrafilter generated by {i}.)

Given that * and $^+$ are boolean homomorphisms, it follows smoothly that $F \in K_{\mathfrak{P}}$. There is hardly any need to check that [ ] can be based on F. □

The above proof also establishes

II.82. THEOREM. The logic determined by the class of all ordering functions is $\mathcal{P}$.[22)]

II.4.3. Extensions of $\mathcal{P}$

II.83. PROPOSITION. Let $<$ be an ordering function for $\mathcal{W}$. Suppose [ ] is an interpretation over $\mathcal{W}$ based on $<$ which verifies MP. Then $<$ can be based on a faithful ordering function for $\mathcal{W}$.

PROOF. Define a new ordering function $<'$ for $\mathcal{W}$ as follows. Let $\mathcal{D}_i' = \mathcal{D}_i \cup \{i\}$ for every $i \in \mathcal{W}$. And set

$$k <_i' j \text{ iff } k <_i j \text{ and } j \neq i$$

Note that $<'$ is transitive, irreflexive and faithful. (And note that even if $<$ is connected or almost-connected, $<'$ need be neither.)

We must show that $i \in [\phi \to \psi]$ iff

$\forall j \in [\phi] \cap \mathcal{D}_i' \; \exists k \in [\phi] \cap \mathcal{D}_i'(k \leq_i' j \;\&\; \forall l \in [\phi] \cap \mathcal{D}_i'(l \leq_i' k \supset l \in [\psi]))$

Suppose $i \in [\phi \to \psi]$. Let j be any world in $[\phi] \cap \mathcal{D}_i'$. Two possibilities obtain.

(a) $j = i$. Since [ ] verifies MP, $i \in [\phi]$, and $i \in [\phi \to \psi]$, we may assume that $i \in [\psi]$. Because there is no $k <_i i$, this suffices.

(b) $j \neq i$. Then we have $\forall k((k \in \mathcal{D}_i' \;\&\; k \leq_i' j) \supset (k \in \mathcal{D}_i \;\&\; k \leq_i j))$. Since $\exists k \in [\phi] \cap \mathcal{D}_i(k \leq_i j \;\&\; \forall l \in [\phi] \cap \mathcal{D}_i(l \leq_i k \supset l \in [\psi]))$, this means there is nothing to prove here.

Conversely, suppose $i \notin [\phi \to \psi]$. Then there is some $j_o \in \mathcal{D}_i \cap [\phi]$ such that

$\forall k \in [\phi] \cap \mathcal{D}_i(k \leq_i j_o \supset \exists l \in [\phi] \cap \mathcal{D}_i(l \leq_i k \;\&\; l \notin [\psi]))$. Again we distinguish two possibilities.

(a)' not $i \leq_i j_o$, or $i \notin [\phi]$, or $i \notin [\psi]$. None of these poses any problems.

(b)' $i \leq_i j_o$, $i \in [\phi]$ and $[\psi]$. In this case there will be some $j_1 \in [\phi] \cap \mathcal{D}_i$ such that $j_1 \leq_i i$ and $j_1 \notin [\psi]$. Because $\leq$ is transitive, it holds that

$\forall k \in [\phi] \cap \mathcal{D}_i(k \leq_i j_1 \supset \exists l \in [\phi] \cap \mathcal{D}_i(l \leq_i k \;\&\; l \notin [\psi]))$.

Note that if $k \leq_i j_1$ then $k \leq_i' j_1$. So we see

$\exists j \in [\phi] \cap \mathcal{D}_i' \; \forall k \in [\phi] \cap \mathcal{D}_i'(k \leq_i' j \supset \exists l \in [\phi] \cap \mathcal{D}_i'(l \leq_i' k \;\&\; l \notin [\psi]))$.

□

In a similar fashion it can be shown that the following holds.

II.84. PROPOSITION. Let $<$ be a faithful ordering for $\mathcal{W}$. Suppose [ ] is an interpretation over $\mathcal{W}$ based on $<$ which verifies CS. Then [ ] can be based on a centered ordering function for $\mathcal{W}$.

PROOF. Define a new ordering function $<'$ for $\mathcal{W}$ by stipulating that

$k <'_i j$ iff (a) $k <_i j$; or (b) $k = i$, $j \neq i$, and $j \in \mathcal{D}_i$.

It is left to the reader to check that everything works out well. (Note in passing that if $<$ is almost-connected, $<'$ will be so too.)

□

II.85. COROLLARIES.

(i) The logic determined by the class of faithful ordering functions is $\mathfrak{P}$ + MP.

(ii) The logic determined by the class of centered ordering functions is $\mathfrak{P}$ + MP + CS.

PROOF. (i) By inserting proposition 83 at the right place in the proof of theorem 81 , this proof can be extended to a proof showing that $\mathfrak{P}$ + MP is complete and determined by the class of faithful ordering functions.

(ii) Analogous. □

If you are wondering why I left out a discussion of $\mathfrak{P}$ + CS, then recall that every ordering function validating $\mathfrak{P}$ + CS also validates MP. ($\mathfrak{P}$ + CS is incomplete with respect to classes of ordering functions so to speak. Cf. section 3.2.)

Let us now turn to $\mathfrak{P}$ + ASP = $\mathfrak{V}$. That $\mathfrak{V}$ is the logic determined by the class of all almost-connected ordering functions was first proved by Lewis (1973). See Krabbe (1978) for a correction. I will give here a new proof of this result.

II.86. LEMMA. The following arguments belong to $\mathfrak{V}$.

(i) $(\phi \vee \psi) \to \sim\phi \;/\; \psi \to \sim\phi$;

(ii) $(\phi \vee \psi) \to \sim\phi,\ (\psi \vee \chi) \to \sim\psi \;/\; (\phi \vee \chi) \to \sim\phi$;

(iii) $(\phi \vee \psi) \to \sim\phi,\ \sim((\psi \vee \chi) \to \sim\chi) \;/\; (\phi \vee \chi) \to \sim\phi$;

(iv) $\phi \to \psi,\ \sim((\phi \vee \psi) \to \sim\chi) \;/\; \chi \to (\sim\phi \vee \psi)$.

PROOF. There are straightforward semantic proofs showing that the arguments mentioned under (i) and (ii) are valid within the class of all ordering functions. Since we already know that all arguments valid within this class belong to $\mathfrak{P}$, we may rest assured that these arguments also belong to $\mathfrak{V}$.

For (iv), you are referred to the proof of proposition 80. There you can find a derivation showing that all arguments of the form $\phi \to \psi,\ (\phi \vee \chi) \to \chi \;/\; \chi \to (\sim\phi \vee \psi)$ belong to $\mathfrak{P}$. By replacing the application of ASC in this derivation by an application of ASP, we get a derivation showing that all arguments of the form $\phi \to \psi,\ \sim((\phi \vee \psi) \to \sim\chi) \;/\; \chi \to (\sim\phi \vee \psi)$ belong to $\mathfrak{V}$.

(iii) is presumably most quickly proved by first giving semantic proofs showing that all arguments of the form

$$(\phi \vee \psi) \to \sim\phi \;/\; (\phi \vee \psi \vee \chi) \to \sim\phi \qquad (*)$$

and all arguments of the form

$$\sim((\psi \vee \chi) \to \sim\chi) \;/\; \sim((\phi \vee \psi \vee \chi) \to \sim(\phi \vee \chi)) \qquad (**)$$

belong to $\mathfrak{P}$. We then have

$$\dfrac{\dfrac{(\phi \vee \chi) \to \sim\phi}{(\phi \vee \psi \vee \chi) \to \sim\phi}\,(*) \qquad \dfrac{\sim((\psi \vee \chi) \to \sim\chi)}{\sim((\phi \vee \psi \vee \chi) \to \sim(\phi \vee \chi))}\,(**)}{(\phi \vee \psi) \to \sim\phi}\,\text{ASP}$$

which shows that all arguments mentioned under (iii) belong to $\mathfrak{P}$ + ASP = $\mathfrak{V}$.

(If you prefer syntactic proofs, here is a derivation of (*).

$$\dfrac{\dfrac{(\phi \vee \psi) \rightarrow \sim\phi}{(\phi \vee \psi) \rightarrow ((\chi \wedge \sim(\phi \vee \psi)) \vee \sim\phi)}\,\text{CW} \qquad \dfrac{\dfrac{\emptyset}{(\chi \wedge \sim(\phi \vee \psi)) \rightarrow (\chi \wedge \sim(\phi \vee \psi))}\,\text{CZ}}{(\chi \wedge \sim(\phi \vee \psi)) \rightarrow ((\chi \wedge \sim(\phi \vee \psi)) \vee \sim\phi)}\,\text{CW}}{\dfrac{(\phi \vee \psi \vee \chi) \rightarrow ((\chi \wedge \sim(\phi \vee \psi)) \vee \sim\phi)}{(\phi \vee \psi \vee \chi) \rightarrow \sim\phi}\,\text{CW}}$$

The next derivation shows that
$(\phi \vee \psi \vee \chi) \rightarrow \sim(\phi \vee \chi) \;/\; (\psi \vee \chi) \rightarrow \sim\chi$ belongs to $\mathfrak{P}$.
Given the rules for negation it follows that also (**) $\in \mathfrak{P}$.

$$\dfrac{\dfrac{(\phi \vee \psi \vee \chi) \rightarrow \sim(\phi \vee \chi)}{(\phi \vee \psi \vee \chi) \rightarrow \sim\chi}\,\text{CW} \qquad \dfrac{\dfrac{(\phi \vee \psi \vee \chi) \rightarrow \sim(\phi \vee \chi) \qquad \dfrac{\emptyset}{(\phi \vee \psi \vee \chi) \rightarrow (\phi \vee \psi \vee \chi)}\,\text{CI}}{(\phi \vee \psi \vee \chi) \rightarrow (\sim(\phi \vee \chi) \wedge (\phi \vee \psi \vee \chi))}\,\text{CC}}{(\phi \vee \psi \vee \chi) \rightarrow (\psi \vee \chi)}\,\text{CW}}{(\psi \vee \chi) \rightarrow \sim\chi}\,\text{ASC}$$

) □

II.87. THEOREM. Let S be a separative selection function for $\mathcal{W}$. Suppose [ ] is a standard interpretation over $\mathcal{W}$ based on S which verifies $\mathfrak{P}$. Then [ ] can be based on an almost-connected ordering function for $\mathcal{W}$.

PROOF. Let $\mathcal{D}_i$ and $p_i(j)$ for every $i,j \in \mathcal{W}$ be defined as in the proof of theorem 80 . This time we consider the ordering function $<$ for $\mathcal{W}$ given by

$k <_i j$ iff $k,j \in \mathcal{D}_i$ and $S_i(p_i(j) \cup p_i(k)) \subset \mathcal{W} \sim p_i(j)$.

Just like before, we will omit the subscript i since no confusion can arise.

Notice that $<$ is irreflexive.

To prove transitivity, we must show that

$$\left.\begin{array}{l} S(p(j) \cup p(k)) \subset \mathcal{W} \sim p(j) \\ S(p(k) \cup p(l)) \subset \mathcal{W} \sim p(k) \end{array}\right\} \Rightarrow S(p(j) \cup p(k)) \subset \mathcal{W} \sim p(j)$$

Given (ii) of the lemma, it is clear that this is indeed the case.

To prove almost-connectedness, we must show that

$$\left.\begin{array}{l} \mathcal{S}(p(j) \cup p(k)) \subset \mathcal{W} \sim p(j) \\ \mathcal{S}(p(k) \cup p(l)) \not\subset \mathcal{W} \sim p(l) \end{array}\right\} \Rightarrow \mathcal{S}(p(j) \cup p(l)) \subset \mathcal{W} \sim p(l)$$

Here we can apply (iii) of the lemma.

It remains to be proved that [ ] can be based on <. To this end we show that $\mathcal{S}([\phi]) \subset [\psi]$ iff $\forall j \in [\phi] \cap \mathcal{D}\ \exists k \in [\phi] \cap \mathcal{D}(k \leq j\ \&\ \forall l \in [\phi] \cap \mathcal{D}(l \leq k \supset l \in [\psi]))$.

Suppose $\mathcal{S}([\phi]) \not\subset [\psi]$. Consider any $j \in \mathcal{S}([\phi]) \sim [\psi]$. Since [ ] verifies $\mathfrak{v}$, $j \in [\phi]$. Also, $j \in \mathcal{D}$. Moreover, if $k < j$ then $k \notin [\phi]$. (For, if $k < j$ then $\mathcal{S}(p(k) \cup p(j)) \subset \mathcal{W} \sim p(j)$; by (i) of the lemma this implies that $\mathcal{S}(p(k)) \subset \mathcal{W} \sim p(j)$; since $k \in \mathcal{S}(p(k))$ and $p(j) = [\phi]$, it follows that $k \notin [\phi]$.)

So we see that

$\exists j \in \mathcal{D} \cap [\phi]\ \forall k \in \mathcal{D} \cap [\phi]\ (k \leq j \supset \exists l \in \mathcal{D} \cap [\phi]\ (l \leq k\ \&\ l \notin [\psi]))$.

To prove the converse, assume that $\mathcal{S}([\phi]) \subset [\psi]$. Consider any $j \in \mathcal{D} \cap [\phi]$. It must be shown that

$\exists k \in \mathcal{D} \cap [\phi](k \leq j\ \&\ \forall l \in \mathcal{D} \cap [\phi](l \leq k \supset l \in [\psi]))$.

As in the proof of theorem 80 , we distinguish two cases:

(a) $\forall l \in \mathcal{D} \cap [\phi](l \leq j \supset l \in [\psi])$. In this case there is nothing to prove.

(b) $\exists l_o \in \mathcal{D} \cap [\phi](l_o \leq j\ \&\ l_o \notin [\psi])$. Since $l_o \in \mathcal{D}$, $l_o \in \mathcal{S}([\chi])$ for some χ.

Claim: $\emptyset \neq \mathcal{S}([\phi \vee \chi]) \subset \mathcal{W} \sim [\chi]$.

Proof of the claim:

Note that $(\phi \vee \chi) \rightarrow \sim\chi$, $(\phi \vee \chi) \rightarrow \chi\ /\ \chi \rightarrow \sim\chi$ is an instance of ASC. So if $\mathcal{S}([\phi \vee \chi])$ would be empty $\mathcal{S}([\chi])$ would be empty, too, which contradicts the assumption that $l_o \in \mathcal{S}([\chi])$. Secondly, if $\mathcal{S}([\phi \vee \chi]) \not\subset \mathcal{W} \sim [\phi]$, we would have that $i \in [\phi \rightarrow \psi]$ and $i \in [\sim((\phi \vee \chi) \rightarrow \sim\chi)]$; then by (iv) of the lemma it would follow that $i \in [\chi \rightarrow (\sim\phi \vee \psi)]$, which contradicts the assumption that $l_o \in \mathcal{S}([\chi])$, $l_o \in [\phi]$ and $l_o \notin [\psi]$. Now consider any $k_o \in \mathcal{S}([\phi \vee \chi])$. Clearly $k_o < l_o$. Furthermore, since $\phi \rightarrow \psi / (\phi \vee \chi) \rightarrow (\psi \vee \chi)$ belongs to $\mathfrak{v}$ and $k_o \notin [\chi]$ we have that $k_o \in [\psi]$. Finally, there is no $l < k_o$ such that $l \in [\phi]$. So trivially

$$\forall l \in \mathcal{D} \cap [\phi](l \leq k_o \supset l \in [\psi]).$$

□

II.88. THEOREM.

(i) $\mathfrak{V}$ is complete.

(ii) The logic determined by the class of all almost-connected ordering functions is $\mathfrak{V}$.

PROOF. By replacing the application of theorem 80 in the proof of theorem 81 by an application of the theorem above, we get a proof of (i). This proof also establishes (ii). □

Perhaps you had expected instead of theorem 87 , a saying that every interpretation which is based on an ordering function and which verifies ASP can be based on an almost-connected ordering function. In other words, why did I not treat $\mathfrak{V}$ as I treated $\mathfrak{P}$ + MP and $\mathfrak{P}$ + MP + CS?

Also, having noticed the similarities between the proofs of theorem 80 and theorem 87 , you may have wondered why I did not reduce them to the same denominator. Can the definition of $<$ in the proof of theorem 80 not be replaced by a less stringent one - one that does not only work for $\mathfrak{P}$, but also for $\mathfrak{P}$ + ASP and perhaps even for other extensions of $\mathfrak{P}$? After all, when we were dealing with questions of interchangeability, we had no problem defining $<$ in such a way that it did not only work for 'plain' partial ordering functions but also for faithful, centered, universal almost-connected and connected ordering functions (see theorem 71) and the subsequent discussion). Why then should there be any problem here?

I am afraid I do not have a satisfactory answer to these questions. It might very well be that the proposition in question holds, but the problem is that I would not know how to prove it except for the special case in which $<$ satisfies the Limit Assumption. And for our purposes this is too weak a result. Likewise, it could be that the definition of $<$ in theorem 80 can be replaced a more fruitful one. I am beginning to believe, however, that there is not very much to be gained.

II.89. THEOREM.

(i) The logic determined by the class of faithful and almost-connected ordering functions is $\mathfrak{V}$ + MP.

(ii) The logic determined by the class of centered and almost-connected ordering functions is $\mathfrak{V}$ + MP + CS.

PROOF.

(i) We cannot apply proposition 83 here (as we already noted parenthetically when we were proving it). Neither is there a guarantee that the ordering function figuring in the proof of theorem 87 will be faithful if the interpretation concerned verifies MP. Still, it is easy to see that if this interpretation happens to verify MP, then there cannot be any $i,k \in \mathcal{D}_i$ such that $k <_i i$. We cannot be sure, however, that i will always be a member of $\mathcal{D}_i$.

It is not difficult to remedy this defect. We adapt the proof of theorem 87 as follows. Set $j \in \mathcal{D}_i$ iff (i) $j = i$; (ii) $j \in \mathcal{D}_i([\phi])$ for some ϕ. Let $p_i(j)$ for every $j \in \mathcal{D}_i$ such that $j \neq i$ be defined as before, and take $p_i(i) = \mathcal{W}$. The definition of < can remain as it is. And so can the remainder of the proof once it has been noted that this time < *is* faithful, and that in virtue of the fact that [ ] satisfies MP, we still have that if $k <_i j$ and $[\phi] \in p_i(j)$, then $k \notin [\phi]$.

(ii) I already announced parenthetically that we would apply proposition 84 here. □

II.90. THEOREM. The logic determined by the class of all connected ordering functions is $\mathfrak{V}$ + CEM.

PROOF. It suffices to show that any standard interpretation that is based on a separative selection function S and that verifies $\mathfrak{V}$ + CEM can be based on a connected ordering function for the domain $\mathcal{W}$ concerned.

Let < be defined as in the proof of theorem 87. We already know that < is irreflexive, transitive, almost connected and that [ ] can be based on <. Call $j,k \in \mathcal{W}$ equivalent iff for every ϕ, $j \in [\phi]$ iff $k \in [\phi]$. Claim: for every $j,k \in \mathcal{W}$, either $j < k$, or $k < j$, or j is equivalent to k.

Proof of the claim. Suppose neither $j < k$ nor $k < j$. Then $\mathcal{S}(p(j) \cup p(k)) \not\subset \mathcal{W} \sim p(k)$. Since $\mathcal{S}$ verifies CEM this means that $\mathcal{S}(p(j) \cup p(k)) \subset p(k)$. Likewise we see that $\mathcal{S}(p(j) \cup p(k)) \subset p(j)$. By applying ASC two times, we see that $\mathcal{S}(p(j)) \subset p(k)$ and $\mathcal{S}(p(k)) \subset p(j)$. Since CE belongs to $\mathfrak{V}$ + CEM it follows that $\mathcal{S}(p(j)) \subset [\phi]$ iff $\mathcal{S}(p(k)) \subset [\phi]$. Now suppose $j \in [\phi]$. Using CEM, we see that $\mathcal{S}(p(j)) \subset [\phi]$. Hence $\mathcal{S}(p(k)) \subset [\phi]$, and therefore $k \in [\phi]$. Obviously, the converse holds, too. It is clear, then, that k and j are equivalent.

Now let $<_i'$ for every i be any irreflexive, transitive and connected extension of $<_i$. Note that if $k <_i' j$ then either $k < j$, or k is equivalent to j. Using this it is easy to check that [ ] can be based on $<'$. □

The case of $\mathfrak{V}$ + CEM + MP is now straightforward to say nothing of $\mathfrak{V}$ + CEM + MP + CS.

The construction used in the proof of theorem 87 also enables us to get to grips with the Limit Assumption.

II.91. THEOREM. Let $\mathcal{S}$ be a separative selection function for $\mathcal{W}$. Suppose [ ] is a standard interpretation over $\mathcal{W}$ based on $\mathcal{S}$ which verifies $\mathfrak{V}$ + LIM. Then [ ] can be based on an almost-connected ordering function satisfying the Limit Assumption.

PROOF. Let < be defined as in the proof of theorem 87. The only thing we have to show is that < satisfies the Limit Assumption. Assume for contradiction that < does not do so. Then there is for some i an infinite sequence of worlds $j_0, \ldots, j_n, \ldots$ such that for all $n \in \omega$, $j_{n+1} <_i j_n$. In view of the definition of < it follows that there are $\phi_0, \ldots, \phi_n, \ldots$ such that for every $n \in \omega$,

$$i \in [(\phi_n \vee \phi_{n+1}) \rightarrow \sim\phi_n].$$

By induction on n, it follows straightforwardly that for every $n \in \omega$

$i \in [(\phi_0 \vee \ldots \vee \phi_n \vee \phi_{n+1}) \to \sim(\phi_1 \vee \ldots \vee \phi_n)$. (a)

(To prove the induction step, you have to apply (*) from lemma 86 twice.) Also, it is clear that for every $n \in \omega$,

$i \in [\sim((\phi_0 \vee \ldots \vee \phi_n \vee \phi_{n+1}) \to (\phi_1 \vee \ldots \vee \phi_n)]$. (b)

(Otherwise, we would have for some $n \in \omega$,

$i \in [(\phi_1 \vee \ldots \vee \phi_n \vee \phi_{n+1}) \to \bot]$.

Which would imply that for every $k \leqq n+1$,

$i \in [\phi_k \to \bot]$.

And this contradicts the fact that for every n, $[\phi_n] = p(j)$ for some $j \in \mathcal{D}_i$.)

Since [ ] verifies LIM, it follows from (a) that $i \in [\bot]$. Contradiction. □

By replacing the application of theorem 80 in the proof of theorem 81 by an application of the above proposition we get a proof of

II.92. COROLLARY. $\mathfrak{V}$ + LIM is canonical.

But since $\mathfrak{V}$ + LIM is not compact we cannot apply theorem 19.

II.93. QUESTION. Is $\mathfrak{V}$ + LIM complete?

COMMENT. If every $\mathfrak{V}$ + LIM - consistent set of sentences is extendable to a *maximal* $\mathfrak{V}$ + LIM - consistent set of sentences, then the proof of theorem 19 can without much ado be adjusted. If on the other hand, some $\mathfrak{V}$ + LIM - consistent set of sentences is not so extendable, then $\mathfrak{V}$ + LIM is incomplete.

The question has some wider interest, since so far no example of incomplete canonical logics - either modal logics, or tense logics, or conditional logics - are known. □

The result reported in theorem 92 can without much ado be transferred to extensions of $\mathfrak{V}$ + LIM. $\mathfrak{V}$ + LIM + CEM, $\mathfrak{V}$ + LIM + CEM + MP, $\mathfrak{V}$ + LIM + MP and $\mathfrak{V}$ + LIM + MP + CS are all canonical, but perhaps none of these logics is complete. The same applies to $\mathfrak{P}$ + LIM and the logics $\mathfrak{P}$ + LIM + MP, and

$\mathfrak{P}$ + LIM + MP + CS, though here one has to take the construction of < used in theorem 80 as the starting point of the proof.

One more remark on the logics containing LIM. Let Δ/ϕ be an argument with finitely many premises. Suppose $\Delta/\phi \in \mathfrak{L}$ + LIM where $\mathfrak{L}$ is any of the logics $\mathfrak{P}$, $\mathfrak{P}$ + MP, $\mathfrak{P}$ + MP + CS, $\mathfrak{V}$, $\mathfrak{V}$ + MP + CS, $\mathfrak{V}$ + CEM, $\mathfrak{V}$ + CEM + MP. Then we have

$\Delta/\phi \in \mathfrak{L}$ + LIM iff $\Delta/\phi \in \mathfrak{L}$.

This is so because all of the logics $\mathfrak{L}$ concerned have the following property.

II.94. THEOREM. Let Δ be a finite $\mathfrak{L}$-consistent set of sentences. Then there is a finite set of possible worlds W, and a standard compositional interpretation [ ] over W, such that (i) [ ] verifies $\mathfrak{L}$, (ii) for some $i \in W$ it holds that $i \in \phi$ for every $\phi \in \Delta$.

PROOF. I will not prove this as it follows immediately from a much more general result in Lewis (1974). □

Not only does it immediately follow from this theorem that all the logics $\mathfrak{L}$ concerned are decidable, but we can also take the finite sets W, together with the standard compositional interpretations [ ] verifying $\mathfrak{L}$ and satisfying Δ, as the starting point for the construction of *finite* sets W', serving as the domain of some suitable ordering function < on which [ ] can be based and which validates $\mathfrak{L}$. Since these ordering functions trivially satisfy the Limit Assumption, the result mentioned follows smoothly.

The only property of ordering functions which we have not yet discussed is Universality. The logical impact of this property is well known, and will not be discussed at any length here. Let $\mathfrak{L}$ be any of the logics figuring in the above. Suppose $\mathfrak{L}$ is determined by the class K of ordering functions. Let K' be the class of all universal ordering functions in K. Then the logic determined by K' is the weakest extension of $\mathfrak{L}$ closed under the following rules

$\Box\phi/\phi$, $\Box\phi/\Box\Box\phi$, $\phi/\Box\Diamond\phi$

Main ingredients of the proof: Let < be any ordering function for $\mathcal{W}$, given $\mathcal{D}$. The argument forms concerned are valid on < iff for any $i,j \in \mathcal{W}$, (i) $i \in \mathcal{D}_i$, (ii) if $j \in \mathcal{D}_i$ then $\mathcal{D}_j \subset \mathcal{D}_i$, (iii) if $j \in \mathcal{D}_i$ then $i \in \mathcal{D}_j$. Furthermore, it is not difficult to show that if Δ is satisfiable on some ordering function with the properties (i), (ii), and (iii), Δ is also satisfiable on some universal ordering function. (If all sentences of Δ happen to be true at i, then take $\mathcal{D}_i$ as the new domain, and restrict <, $\mathcal{D}$ and [ ] to $\mathcal{D}_i$.)

NOTES TO PART II

1. In doing so we also meet the wishes of cognitive psychologists like Philip Johnson-Laird (1985), who argues that the logical theories for conditionals developed within the framework of possible-worlds semantics cannot serve as a basis for a *cognitive* theory of conditionals because, as he puts it 'the set of possible worlds goes far beyond what can fit into an individual's mind ... Each possible world exceeds what any individual can apprehend ... If Stalnaker, Lewis or their colleagues have defined the truth conditions of conditionals, then no one can ever grasp them, and *a fortiori* no one can ever properly evaluate any conditional.' (Additional note: By lumping all possible worlds semanticists together, Johnson-Laird is not doing justice to in particular Stalnaker.)

2. It is in particular this aspect of possible worlds semantics that has been much under attack recently. Here data semantics (see part III) may serve as an example.

3. An exception, or rather an attempt to develop an exception, can be found in chapter 2 of Nute (1980). But in the meantime Nute has concluded that this proposal has 'very serious difficulties', and 'appears to be a dead end'. See Nute (1984: 416).

4. Consider the *relation* $f_{\sim}$ defined by: $\langle p,q \rangle \in f_{\sim}$ iff for some ϕ, $p = [\phi]$ and $q = [\sim\phi]$. If $f_{\sim}$ is not a *function*, then there are ϕ and ψ such that $[\phi] = [\psi]$ and $[\sim\phi] \neq [\sim\psi]$. But this immediately shows that the principle of replacement does not hold. That this principle neither holds if $f_{\wedge}$ or $f_{\rightarrow}$

are no functions, can be shown in the same manner. That it does hold if both $f_{\sim}$, $f_{\wedge}$ and $f_{\rightarrow}$ are functions can be, as well as proved by showing with induction on the complexity of χ that whenever $[\phi] = [\psi]$, the replacement of an occurrence of ϕ in χ by an occurrence of ψ always yields a sentence χ' such that $[\chi] = [\chi']$.

5. The exceptions being R. Routley and R.K. Meyer (see Routley and Meyer (1973). However, their theory admits inconsistent as well as incomplete possible worlds, things that are not possible worlds in our sense of the word.

6. Henceforth reference to L is suppressed.

7. Here we depart from standard practice. A logic that is complete in our sense of the word is usually called *generally complete*. And a *complete* logic in the usual sense of the word would be a logic $\mathcal{L}$ for which there is a class K of connectors such that $\{\phi \mid \emptyset/\phi \in \mathcal{L}\} = \{\phi \mid \phi \text{ is valid within } K\}$.

8. Let F be a connector with domain W, and G be a connector with domain V. G is *isomorphic* to F iff there is a function c, mapping W one-to-one onto V such that for any $p,q \subset W$ and $i \in W$

$$i \in F(p,q) \text{ iff } c(i) \in G(\bar{c}(p),\bar{c}(q)).$$

(Here $\bar{c}$ is the function from pow W onto pow V given by

$$\bar{c}(p) = \{j \in V \mid j = c(i) \text{ for some } i \in p\})$$

Note that the isomorphism relation is an equavalence relation. Moreover for isomorphic F and G the following holds: let I be an atomic interpretation over W, and J be an atomic interpretation over V. Suppose that for all atomic ϕ, $J(\phi) = \bar{c}(I(\phi))$. Then the interpretation [ ] conforming to I and based on F, and the interpretation $[\![\ ]\!]$ conforming to J and based on G are likewise related, i.e. for every ϕ, $[\![\phi]\!] = \bar{c}([\phi])$. This means that every Δ satisfiable by F will be satisfiable by G, and *vice versa*. So if $F \in K$, $G \in K_{(\mathcal{L}_K)}$.

9. The examples given show in fact that there are logics with the property that there is no class K of connectors such that $\{\phi \mid \phi$ is valid within $K\} = \{\phi \mid \emptyset/\phi \in \mathcal{L}\}$, which is stronger than is needed for our purposes here.

10. Consequence functions were first introduced in Chellas (1975), albeit under a different name.

11. Lewis (1973: 58) credits John Vickers with the introduction of selection functions.

12. The assumption that any world j is accessible from any world i (formally: $\mathcal{D}_i = W$ for any i) yields the theory of strict implication that we discussed in section I.2.1.1.

13. Note that not all quantifiers between *by far the most* and *all* satisfy (ii).For example, consider the quantifier $\mathcal{Q}$ defined by

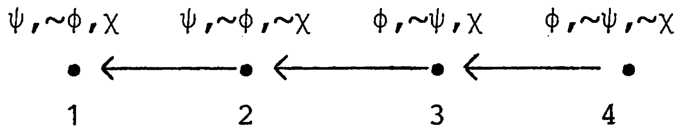
$$\mathcal{Q}_W(p,q) \text{ iff } \begin{cases} |p \sim q| = 0, \text{ when p is finite} \\ |p \sim q| \text{ is finite and even, when p is infinite} \end{cases}$$

Obviously, this quantifier does not satisfy the condition that $\mathcal{Q}_W(p,r)$ if $\mathcal{Q}_W(p,q)$ and $q \subset r$. (This example was brought to my notice by Peter van Emde Boas.)

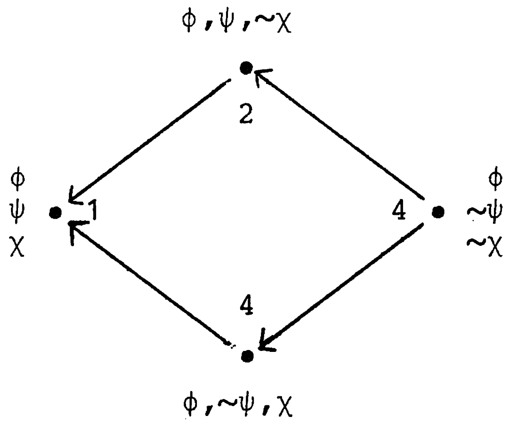
14. Given Aug and Cut this form of the principle of Modus Ponens is equivalent to the one introduced in section 2.1.

15. This is how Lewis described the Limit Assumption in his (1981). In Lewis (1973) a much weaker version relative to interpretations is discussed which says that for every sentence ϕ, such that $[\phi] \cap \mathcal{D}_i \neq \emptyset$ there is some $[\phi]$-closed world to i.

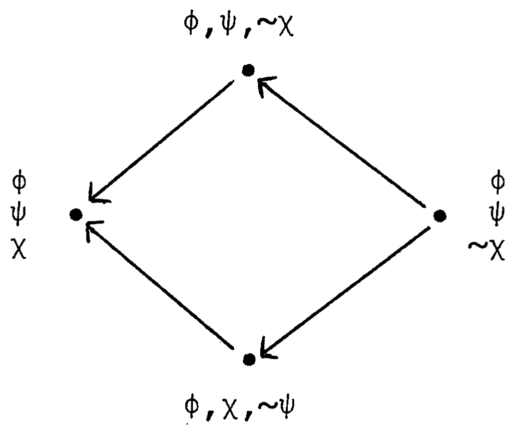
16. Counterexamples to AD, CC, and ASP are pictured below. An arrow is drawn from k to j *if and only if* $j <_1 k$



Note that $1 \in [\phi \rightarrow \chi]$, $1 \in [\psi \rightarrow \chi]$, $1 \notin [(\phi \vee \psi) \rightarrow \chi]$.



Note that $1 \in [\phi \rightarrow \chi]$, $1 \in [\phi \rightarrow \psi]$, $1 \notin [\phi \rightarrow (\psi \wedge \chi)]$.



Note that $1 \in [\phi \rightarrow \chi]$, $1 \in [\phi \rightarrow \psi]$, $1 \notin [(\phi \wedge \psi) \rightarrow \chi]$. These are the simplest counterexamples I have been able to find.

17. John Pollock has tried on various occasions, most recently in Pollock (1981).For a rejoinder, see Nute (1984: 435).

18. This result is mentioned in Veltman (1976). Its importance, however, was first noticed by Kratzer (1981).

19. The result is due to Lewis (1981).

20. See Lewis (1979).

21. Johan van Benthem suggested a different proof of this theorem which starts from the following definitions:
$\mathcal{D}_i = \{j \subseteq \mathcal{W} \mid j \in S_i(p) \text{ for some } p\}$
$j \leqq_i k$ iff $j,k \in \mathcal{D}_i$ and $S_i\{k,j\} \subset \{j\}$

22. John Burgess (1979) gave a proof of the somewhat weaker theorem that the set of sentences valid on all ordering functions coincides with $\{\phi \mid \emptyset/\varphi \in \mathcal{P}\}$.
I learnt a lot from his proof.

PART III

D A T A   S E M A N T I C S

III.1. SEMANTIC STABILITY AND INSTABILITY

The semantic system developed here differs in various respects from the kind of systems developed in part II. These differences have largely to do with the role which information plays. In part II we concluded that background information plays an important role in the evaluation of counterfactuals. Premise semantics can be seen as a first attempt to make this role explicit. In data semantics it becomes even more important, since whereas in premise semantics the evaluation of just some kinds of sentences is made dependent on background information, in data semantics this is extended to all kinds of sentences. This more radical approach derives from the idea that the meaning of many kinds of expressions is deeply bound up with the restricted knowledge which we have of the world in which we live. We use language not only to say what we know about the world but also to express our ignorance of it.

III.1.1. Information models

As we already saw in the introduction (I.2.1.2), the central concept in data semantics is not truth *simpliciter* but truth *on the basis of the available evidence*. Consequently, our principal concern will be to sort out what it means for a sentence to have this property.

Following usual logical practice, I shall not deal with this directly but introduce a formal language *L*, the sentences of which will serve as 'translations' of English sentences. *L* is given by:

(i) a vocabulary consisting of countably many atomic sentences, two parentheses, three one-place operators $\sim$, *must*, and *may* and three two-place operators $\wedge$, $\vee$, and $\rightarrow$.

(ii) the formation rules that one would expect for a language with such a vocabulary.

As usual the operators $\sim$, $\wedge$ and $\vee$ are meant as formal counterparts of negation, conjunction and disjunction respectively. If ϕ and ψ are formal translations of the English sentences ϕ' and ψ', then $\phi \rightarrow \psi$ is meant to be a formal translation of the *indicative* conditional with antecedent ϕ' and consequent ψ'. The operation *may* represents the expression 'it may be the case that', and the operator *must* the expression 'it must be the case that'. It will appear that the semantic and pragmatic properties of indicative conditionals are closely bound up with the properties of these modal expressions, which is why I have included them in *L*.

Up to section 4.2. we will be primarily concerned with sentences which are in a sense simple. To be more precise, we will be concerned with sentences of which the depth does not exceed 1, where the depth of ϕ, $d(\phi)$ is determined as follows:

if ϕ is atomic then $d(\phi) = 0$;
if $\phi = \sim\psi$ then $d(\phi) = d(\psi)$;
if $\phi = \psi \vee \chi$ or $\phi = \psi \wedge \chi$ then $d(\phi) = \text{maximum}(d(\psi), d(\chi))$;
if $\phi = \textit{may}\ \phi$ or $\phi = \textit{must}\ \phi$ then $d(\phi) = d(\psi) + 1$;
if $\phi = \psi \rightarrow \chi$ then $d(\phi) = \text{maximum}(d(\psi), d(\chi)) + 1$.

In presenting the semantics of *L* we shall again follow usual logical practice and first specify the admissable models for *L*.

III.1. DEFINITION. An *information model* (for L) is a triple $\langle S, \leq, V \rangle$ with the following properties

(i) $S \neq \emptyset$

(ii) $\leq$ is a partial ordering of S

(iii) V is a function with domain S; for each $s \in S$, V_s is a partial function assigning at most one of the values 1 or 0 to the atomic sentences of L.
If $s \leq s'$, $V_s \subset V_{s'}$.

An information model is *closed* iff in addition to (i), (ii) and (iii) it has the following property.

(iv) Each maximal chain in $\langle S, \leq \rangle$ contains a maximal element. If s is a maximal element of $\langle S, \leq \rangle$, V_s is total. □

The basic entities of an information model, the elements of S, are called *(possible) information states*: the speakers of the language L - one speaker at different times, or different speakers at the same time - can have different information about a particular state of affairs.

For our purposes all there is to know about any information state is covered by the relation $\leq$ and the function V. V tells for each atomic sentence ϕ and each information state s whether ϕ is true on the basis of the evidence available at s, in which case $V_s(\phi) = 1$, or whether ϕ is false on that basis, in which case $V_s(\phi) = 0$, or whether the evidence available at s does not allow any definite conclusion about the truth value of ϕ, in which case $V_s(\phi)$ is undefined. The relation $\leq$ determines the position of each information state among the others. In this connection it is particularly important to know, given the evidence at a certain information state, what the outcome of any further investigations might be. Whenever $s \leq s'$, we say that *it is possible for* s *to grow into* s'. So understood, it will be clear why $\leq$ is taken to be a partial order.

The requirement that $V_s \subset V_{s'}$, if $s \leq s'$ constrains the semantic properties of atomic sentences considerably: once an atomic sentence ϕ has turned out to be true (or false) on the basis of the evidence, it will remain true (or false), whatever additional data may come to light. As we shall see in the next section, not every sentence of L is *stable* in this sense. Notice that it may very well be that s < s' while $V_s = V_{s'}$: accumulation of evidence need not necessarily mean that more atomic sentences get a definite truth value. (Suppose it is possible for s to grow into an information state where both the atomic sentence ϕ and the atomic sentence ψ are true. It may very well be that this possibility is excluded once s has grown into s'. That does not mean, however, that it must be clear at s' which of the atomic sentences ϕ and ψ is false.)

It remains to explain condition (iv). Consider any subset S' of S. S' is called a *chain* (in $\langle S, \leq \rangle$) iff the restriction of $\leq$ to S' is a linear order. Think of a chain as a sequence of increasingly rich information states. A chain S' is *maximal* iff every chain S'' containing S' is identical to S'. Think of a maximal chain as a sequence of successive information states which is in no way extendable. Now, (iv) says that any such sequence contains a state s' where the information is rich*est*: if S' is a maximal chain then there is some s' $\in S'$ such that for no s $\in S$, s' < s. Moreover, in such a final information state s', the information is, in fact, *complete*: $V_{s'}$ assigns a definite truth value to every atomic sentence. In other words, (iv) excludes the possibility of there being any sequence of successive information states that does not ultimately end in an information state that is complete. Every incomplete information state can grow into a complete information state - in principle, that is, not necessarily in practice.

In the following we will mostly be concerned with closed information models. This is partly for technical convenience. It will appear that as far as logic is concerned everything turns out the same if (iv) is replaced by the much weaker assumption

(iv)' For every s $\in S$ and every atomic sentence ϕ there is some s' $\geq$ s such that $V_{s'}(\phi)$ is defined.

But many results are more easily proved for information models satisfying the stronger condition (iv) than for information models satisfying only (iv)'. (Note that for *finite* models (iv) and (iv)' are equivalent.)

The other reason for restricting ourselves to closed information models is that (iv)' seems natural. What it amounts to is just this: the question whether a given atomic sentence holds or not always *has* an answer. (Once again: whether it will always be possible to really *find* this answer is an altogether different question.) So (iv)' works as a constraint of meaningfulness. And in the context of propositional logic, where nothing can be said about the reasons why certain atomic sentences might be meaningless, imposing this constraint seems a natural thing to do.

Now, let $M = \langle S, \leqq, V \rangle$ be any closed information model. One of the maximal elements of S plays a special role. At that point, say s_o, the information is not only complete, but also *correct*: the evidence available at s_o comprises exactly what is in fact the case. Since the speakers of the language L cannot but get their data from what is in fact the case, they will always be in an information state - perhaps 'evidential situation' would be a better term [2)] - that can grow into s_o. However, as long as their data are incomplete, they do not exactly know what holds at s_o and what does not. That is where the information states that cannot grow into s_o come in: a speaker may at a given point have to reckon with the possibility that further investigations bring him in such an information state even if this does not in fact happen. (Of course, as time goes on, not only may the available information change, but also the facts of the matter. The latter possibility is neglected here. Think of it this way: even if the facts have changed, the investigations under consideration continue, but in such a case they pertain to the past instead of to the present.)

Definition 1 leaves many questions unanswered. For one thing, whenever $V_s(\phi) = 1$ for a given atomic sentence ϕ and information state s we say 'ϕ is true on the basis of the evidence available at s', but the model $\langle S, \leqq, V \rangle$ does not give us any clue as to what this evidence consists in.

Another point is that apart from the requirement that there be sufficiently many complete information states, no constraints at all are placed on the kinds of information states that an information model should contain. The so-called data models introduced in section 4.1. are more satisfactory in these respects. What is here called the evidence available at a given information state, is there identified with a data set, this being a special kind of subset of 'the' set of possible facts. What is here taken as the extension relation '$\leqq$' between information states, there boils down to the subset relation '$\subset$' between data sets. Here, a speaker is supposed to be in a given information state; there he is supposed to be acquainted with the facts that constitute a given data set. Within data models atomic sentences are treated as names of possible facts. The atomic sentence ϕ is true on the basis of a given data set iff this set contains the fact named by ϕ. And ϕ is false on the basis of a given data set iff this set contains a fact that is incompatible with the fact named by ϕ. Thus, in data models the condition that atomic sentences should remain true (or false) once their truth (or falsity) has been established - which we had to stipulate for information models - is automatically fulfilled.

The reason why I mention data models here is because there is one particular feature of information models that cannot be properly explained without reference to data sets. Notice that the information models $<S,\leqq,V>$ are so defined that it may very well occur that for a given atomic sentence ϕ and an information state s the following holds:

(i) for no s' $\geqq$ s, $V_{s'}(\phi) = 0$;

(ii) $V_s(\phi)$ is undefined.

From (i) it follows that $V_{s'}(\phi) = 1$ for every complete s' $\geqq$ s. So it may very well occur that a certain atomic sentence ϕ is not true on the basis of the evidence available at s while on the other hand it is impossible for s to grow into an information state at which ϕ will turn out false. Indeed, s will inevitably grow into an information state at which ϕ is true.

One may wonder whether we should allow this. Wouldn't it be plausible to call ϕ true on the basis of the evidence available at s? Shouldn't we demand that $V_s(\phi) = 1$ if for no $s' \geqq s$, $V_{s'}(\phi) = 0$?

I do not think so. I think it would blur an important distinction - that between *direct* and *indirect* evidence - if one were to maintain that it is solely on the basis of the evidence available at s that the sentence ϕ is true. In the terminology of data sets: speakers in information state s are not directly acquainted with the fact named by ϕ. Their data at best enable them to infer that no fact incompatible with the fact named by ϕ can be added. In other words, their data at best constitute indirect evidence for the truth of ϕ: ϕ must be true, all right, but it may take quite some time until this is definitely shown.

III.1.2. Between assertability and truth

Let $M = \langle S, \leqq, V\rangle$ be any (closed) information model, s an information state in S, and ϕ a sentence. In the sequel,

$$s \models_M \phi$$

abbreviates 'ϕ is true (in M) on the basis of the evidence available at s' and

$$s \;{}_M\!\!=\!\!| \;\phi$$

abbreviates 'ϕ is false (in M) on the basis of the evidence available at s'. When no confusion can arise as to which model M is meant, the subscript 'M' in '$s \;{}_M\!\!=\!\!|\; \phi$' and '$s \models_M \phi$' will be omitted.

The following definition specifies, for any model M, the extension of the relations $\models$ and $=\!\!|$. It applies to arbitrary sentences, but in the course of this study it will appear that for sentences of depth greater than one some modifications are called for.

III.2. DEFINITION. Let $M = \langle S, \leqq, V\rangle$ be any closed information model and s an information state in S.

- If ϕ is atomic, then
  $s \models \phi$ iff $V_s(\phi) = 1$
  $s =\!| \ \phi$ iff $V_s(\phi) = 0$
- $s \models \sim\phi$ iff $s =\!| \ \phi$
  $s =\!| \ \sim\phi$ iff $s \models \phi$
- $s \models \mathit{may}\,\phi$ iff for some $s' \geqq s$, $s' \models \phi$
  $s =\!| \ \mathit{may}\,\phi$ iff for no $s' \geqq s$, $s' \models \phi$
- $s \models \mathit{must}\,\phi$ iff for no $s' \geqq s$, $s' =\!| \ \phi$
  $s =\!| \ \mathit{must}\,\phi$ iff for some $s' \geqq s$, $s' =\!| \ \phi$
- $s \models \phi \wedge \psi$ iff $s \models \phi$ and $s \models \psi$
  $s =\!| \ \phi \wedge \psi$ iff $s =\!| \ \phi$ or $s =\!| \ \psi$
- $s \models \phi \vee \psi$ iff $s \models \phi$ or $s \models \psi$
  $s =\!| \ \phi \vee \psi$ iff $s =\!| \ \phi$ and $s =\!| \ \psi$
- $s \models \phi \rightarrow \psi$ iff for no $s' \geqq s$, $s' \models \phi$ and $s' =\!| \ \psi$
  $s =\!| \ \phi \rightarrow \psi$ iff for some $s' \geqq s$, $s' \models \phi$ and $s' =\!| \ \psi$

In discussing this definition I shall often refer to the following information states.

Information state 1. You are presented with two little boxes, box 1 and box 2. The boxes are closed but you know that together they contain three marbles, a blue one, a yellow one, and a red one, and that each box contains at least one of them.

Information state 2. As 1, except that in addition you know that the blue marble is in box 1. Where the other two marbles are remains a secret.

(So we are back where we started: the case of the marbles.)

Negation

I trust that the truth and falsity conditions for sentences of the form $\sim\phi$ do not need any further explanation. It may, however, be illuminating to compare these conditions with a few alternatives. Presumably, it will not be difficult to convince you that the following stipulation would have been completely mistaken.

(*) $s \models \sim\phi$ iff $s \not\models \phi$

If the evidence available in information state 1 does not allow you to conclude that the blue marble is in box 1, this does not mean that it allows you to conclude that the blue marble is not in box 1. Hence, (*) does not generally hold. (*) only holds if the evidence in s happens to be complete but that is a rather exceptional case. [3)]

Readers familiar with Kripke's semantics for intuitionistic logic will be attracted to the following alternative to the account of negation given in definition 2.

(**) $s \models \sim\phi$ iff for no $s' \geq s$, $s \models \phi$

I can hardly imagine that anyone would adhere to (**) and yet agree with the falsity conditions proposed in definition 2. There seem to be no grounds for denying that the following two statements are equivalent.

(i) ϕ is false on the basis of the available evidence;

(ii) the negation of ϕ is true on the basis of the available evidence.

So I would expect the supporters of (**) to completely reject our falsity conditions rather than to reject the equivalence between (i) and (ii). The incorporation of (**) in definition 2, therefore, would almost certainly bring a drastic revision of the entire system along with it.

At this moment, we are not yet in a position to explain in detail why definition 2 offers a better analysis of negation than (**) does. I shall here briefly sketch the relevant argument trusting that the remainder of this chapter will enable you to fill in the details for yourself.

To begin with, it is worth noting that the negation described by (**) is expressible within the framework presented here, albeit not by means of the operator ~. Still, we have

for no $s' \geq s$, $s \models \phi$ iff $s = $ *must* $\sim\phi$

Hence, the easiest way to compare the (**)-negation and the negation of definition 2, is to study the different properties attributed by definition 2 to sentences of the form *must* $\sim\phi$ on the one hand, and sentences of the form $\sim\phi$ on the other. By doing so for different kinds of sentences, you will undoubtedly sooner or later arrive at the conclusion that 'not' has more in common with the operator ~ than with the operator *must* ~. The following cases are decisive: (i) ϕ is a sentence of the form $(\psi \rightarrow \chi)$; (ii) ϕ is a sentence of the form *must* ψ.

May

Suppose you are in information state 1. Somebody says: 'The blue marble may be in box 2.' Would you agree?
Suppose you are in information state 2. Somebody says: 'The blue marble may be in box 2.' Would you still agree?

According to definition 2, your answer to the first question should be 'Yes', and to the second question 'No'. Definition 2 says that a sentence of the form *may* ϕ is true on the basis of the evidence available at a given information state s as long as it it possible for s to grow into an information state s', where, on the basis of the then available evidence, ϕ is true; and that such a sentence is false on the basis of the evidence available at s iff this possibility is excluded. In information state 1 you must still reckon with the possibility that the blue marble will turn out to be in box 2. Therefore the sentence 'The blue marble may be in box 2' is true on the basis of the evidence available there. In information state 2 you do not have to reckon with this possibility anymore. Once you know that the blue marble is in box 1 it is wrong to maintain that it may nevertheless be in box 2. At most you can say that it might have been in box 2.

Unlike atomic sentences, the truth of sentences of the form *may* ϕ need not be stable. They will often be true on the basis of limited evidence only to become false as soon as new evidence becomes available. Once their falsity has been established, however, it has been established for good. In terms of the following definition: sentences of the form *may* ϕ, though in general not T-stable, are at least F-stable.

III.3. DEFINITION. Let ϕ be a sentence.
ϕ is *T-stable* iff for every model $M = \langle S, \leq, V \rangle$ and information state $s \in S$, if $s \models_M \phi$ then $s' \models_M \phi$ for every information state $s' \geq s$;
ϕ is *F-stable* iff for every model $M = \langle S, \leq, V \rangle$ and information state $s \in S$, if $s\ {}_M{=}\!\!| \ \phi$ then $s'\ {}_M{=}\!\!| \ \phi$ for every information state $s' \geq s$;
ϕ is *stable* iff ϕ is both T-stable and F-stable.

The theory of 'may' developed here differs widely from those developed within the framework of possible worlds semantics. (See II.3.4.) It renders the sentence

(a) *The blue marble is in box 1 and it may not be there*

a logical absurdity, just like

(b) *The blue marble is in box 1 and it isn't*

According to all other theories (a) is a pragmatic, rather than a logical absurdity: (a) can be perfectly true although nobody can ever sincerely assert it.

Is there any *empirical* evidence in favour of this claim, that sentences like (a) are pragmatically rather than logically absurd? I do not think so. The only empirical support which it could conceivably get should consist in an informal example which shows that the apparent inconsistency of sentences of the form $\phi \wedge$ *may* $\sim\phi$ can sometimes be cancelled. I am pretty sure, however, that no such example will ever be found. Anyone asserting a sentence like (a) fails to fulfill the conversational maxim of quality as for example Groenendijk & Stokhof (1975) are ready to explain. (Roughly: by asserting the right hand conjunct 'the blue marble may not be in box 1' one indicates that the sentence 'the blue marble is not in box 1' is consistent with everything one believes. But according to the maxim of quality one is not allowed to assert the left hand conjunct if one does not believe that the blue marble is in box 1.) So if there is any example showing that the apparent inconsistency of these sentences can really be cancelled, it must be one in which one indicates (either explicitly or implicitly, but at least in a way clear enough to the hearer) that one is stating something one does not, in fact, believe, but that this is done for a good reason - a reason which can be reconciled with the overall Cooperative Principle. I am afraid that no hearer will ever be found who is able to detect what good reason that might be.

That it is impossible to breach the maxim of quality and yet observe the overall Cooperative Principle has been noticed before.[4)] For example, Gazdar (1979: 46) notices that an implicature arising from the maxim of quality 'differs from those arising from other maxims because it cannot be intelligibly cancelled'. Yet the only conclusion which is usually drawn is that the maxim of quality has a privileged position among the other maxims. Everybody seems to accept, if reluctantly, that the criterion of cancellability offers at best a *sufficient* condition for calling something pragmatic instead of logical.

The one argument I have to offer in favour of the position that sentences of the form $\phi \wedge$ *may* $\sim\phi$ are logically rather than pragmatically absurd is highly theoretical. Consider the following (re)formulation of the

Maxim of Quality: Do not assert a sentence ϕ *unless* ϕ *is true on the basis of the evidence at your disposal.*

Notice that every sentence which owes its pragmatic absurdity simply and solely to the fact that it can never be asserted without violating this maxim is also absurd for semantic reasons - for *data* semantic reasons at least. Hence the question of cancellability need not arise. By doing data semantics instead of the usual truth-conditional semantics, we have so to speak annexed part of what was always called pragmatics. As a consequence the border between logical and pragmatic-but-not-logical inconsistency and that between logical and pragmatic-but-not-logical validity has been redrawn. Actually it seems that now cancellability can serve as a condition which an argument must satisfy in order to be classified as pragmatically but not logically valid.

These considerations are meant to raise an issue rather than to settle it. It could be that within the standard framework better explanations can be given for the fact that the 'apparent' inconsistency of sentences of the form $\phi \wedge$ *may* $\sim\phi$ cannot be cancelled. Or it may turn out that the framework presented here, just like the standard one, will have to allow for uncancellable implicatures. Moreover, even if neither of these possibilities turns out to be the case, the arguments presented above may not on their own be strong enough to decide which framework is preferable. Other things than a neatly drawn line between semantics and pragmatics may have to be taken into account.

Must

I already hinted at the truth condition for the operator *must* near the end of section 1. According to definition 2, a sentence of the form *must* ϕ is true on the basis of the available evidence iff no additional evidence could make ϕ false. Hence, if one keeps on gathering information, ϕ will inevitably sooner or later turn out true. As long as ϕ could yet turn out false, *must* ϕ is false.

It is worth noting that in many cases this analysis renders a sentence of the form *must* ϕ weaker than ϕ itself. If an atomic sentence ϕ is true on the basis of the available evidence, then *must* ϕ is true on that basis as well. But *must* ϕ can be true on the basis of the evidence without ϕ being true on that basis. In the latter case the data constitute at best indirect evidence for ϕ, in the first case direct evidence.

That *must* ϕ is often weaker than ϕ has been noticed by a number of authors. Karttunen (1972: 12) illustrates this with the following examples:

(a) *John must have left*

(b) *John has left*

His informal explanation fits in neatly with my formal analysis:

'Intuitively, (a) makes a weaker claim than (b). In general, one would use (a), the epistemic *must*, only in circumstances where it is not yet an established fact that John has left. In (a), the speaker indicates that he has no first hand evidence about John's departure, and neither has it been reported to him by trustworthy sources. Instead (a) seems to say that the truth of *John has left* in some way logically follows from other facts the speaker knows and some reasonable assumptions that he is willing to entertain. A man who has actually seen John leave or has read about it in the newspaper would not ordinarily assert (a), since he is in the position to make the stronger claim in (b).'

Similar remarks can be found in Groenendijk & Stokhof (1975), Veltman (1976), Kratzer (1977), and Lyons (1977). Still, despite the unanimity on this point no theory has yet been proposed which actually predicts that on many occasions *must* ϕ is a logical consequence of ϕ. Most theories treat *may* and *must* as epistemic modalities and depending on whether the underlying epistemic notion is knowledge or belief *must* ϕ turns out to be either stronger than ϕ or independent of it. (Cf. II.3.4.)

Notice that sentences of the form *must* ϕ are T-stable though they are not in general F-stable. Consider for example the sentence

(c) *Either the yellow or the red marble must be in box 2*

For all you know in information state 1 it may very well be that the blue marble is in box 2, while both the yellow and the red marble are in box 1. Hence it is not the case that either the yellow or the red one must be in box 2. But as soon as you are told that the blue marble is in box 1 this is different. At least one of the marbles is in box 2 and it cannot be the blue one. So it must be the yellow one or the red one.

If

According to definition 2, a sentence of the form $\phi \rightarrow \psi$ is true on the basis of the evidence available at a given information state s iff s cannot grow into an information state s' at which ϕ is true on the basis of available evidence and ψ is false. If, by any chance, further investigations should reveal that ϕ is true they will also reveal that ψ is true. Furthermore it is stated that $\phi \rightarrow \psi$ is false on the basis of the evidence available at a certain information state s iff it is still possible for s to grow into an information state at which ϕ is true and ψ false on the basis of the available evidence.

As a consequence we find that sentences of the form $\phi \rightarrow \psi$ are not in general F-stable. Consider the sentence

(a) *If the yellow marble is in box 1, the red one is in box 2*

Again, the evidence available in information state 1 allows for the possibility that both the yellow and the red marble are in box 1. So on the basis of the limited evidence available there (a) is false: it is not so that if the yellow marble is in box 1, the red one is in box 2. In information state 2, however, (a) is not false anymore. Once you know that the blue marble is in box 1, you can be sure that if the yellow marble happens to be in box 1, the red one will turn out to be in box 2.

Now consider the negation of (a).

(b) *It is not so that if the yellow marble is in box 1, the red one is in box 2*

This sentence is true on the basis of the evidence available at information state 1 - at least if we apply definition 2 to it. Suppose you are in information state 1 and somebody - Mrs. S. - asserts (a): 'If the yellow marble is in box 1, the red marble is in box 2.' Would it be appropriate, then, to reply like this: 'No, you are wrong, it may very well be that both the yellow and the red marble are in box 1. So it is not the case that if the yellow marble is in box 1, the red one is in box 2'?

Such a reply would only under very special circumstances be correct. Only when you know for certain that Mrs. S. is not better informed than yourself, because only then can you be sure that she is mistaken. Certainly, for all *you* know (in information state 1), sentence (a) is false and sentence (b) is true, but sentence (a) is not F-stable and sentence (b) is not T-stable. If by any chance the blue marble should be in box 1 and if Mrs. S. should know this, then what she says is true on the basis of the evidence available to *her*. So perhaps she is better informed than yourself, perhaps she is telling you something about the marbles you did not yet know. Therefore, instead of denying the truth of her statement you'd better ask yet on what evidence it is based.

In normal conversation every statement is meant to convey some new information and only when this new information is incompatible with some T-stable sentence that is true on the basis of the evidence gathered may one raise doubts about it. Like when you are in information state 2 and Mrs. S. says 'Maybe the yellow marble is in box 1 and *if so, the red one is in box 1 too*'. However, even in this case it would be inappropriate to reply with a simple denial: 'No, it *may* very well be that the yellow marble is in box 1 and the red one is in box 2.' Again, such a sentence is not T-stable; it might owe its truth to a lack of information on your part - that is certainly what Mrs. S. will think. So what you will have to reply is something much stronger: 'No, it *cannot* be that the yellow and the red marble are both in box 1. *If the yellow marble is in box 1, the red one isn't.*'

These considerations may help us to understand some of the peculiarities of negated conditionals. For one thing, they explain why a conditional statement $\phi \rightarrow \psi$ is so often refuted with a counterconditional $\phi \rightarrow \sim\psi$ rather than with a negated conditional $\sim(\phi \rightarrow \psi)$. But they do so without thereby equating sentences of the form $\phi \rightarrow \sim\psi$ with sentences of the form $\sim(\phi \rightarrow \psi)$. On the account given here, $\sim(\phi \rightarrow \psi)$ is not logically equivalent to $\phi \rightarrow \sim\psi$, as it would be if $\rightarrow$ behaved as Stalnaker (1968) and Adams (1975) predict. Nor is it equivalent to $\phi \wedge \sim\psi$ as it would be if $\rightarrow$ behaved like material implication. We find that $\sim(\phi \rightarrow \psi)$ is equivalent to *may* $(\phi \wedge \sim\psi)$.

Let ϕ be F-stable and suppose that ϕ is false on the basis of the available evidence. Then according to definition 2, $\phi \rightarrow \psi$ is true on the basis of the evidence for any sentence ψ. Similarly, if ψ is T-stable and true on the basis of the available evidence then $\phi \rightarrow \psi$ is true on the basis of the evidence for any sentence ϕ. In other words, the present treatment of conditionals does not meet the requirement that a sentence of the form $\phi \rightarrow \psi$ should never be true unless the antecedent ϕ is somehow 'relevant' to the consequent ψ. The well-known 'paradoxes' of material implication turn out logically valid. We find, for example, that from a logical point of view, there is nothing wrong with

(a) *The blue marble is in box 1*

∴ *If the blue marble is in box 2, it is in box 1*

If you do find it difficult to accept the validity of this argument, please read the conclusion once more without losing sight of the premise. The argument does not run like

(b) *The blue marble is in box 1*

∴ *If the blue marble had been in box 2, it would have been in box 1*

Or perhaps it helps to compare (a) with

(c) *The blue marble is in box 1*

∴ *The blue marble is in box 1, if it is anywhere at all*

(Anywhere ... then why not try box 2.) If this does not help either, the reader is referred to section 2.3, where I shall argue that (a), though logically valid, is nevertheless pragmatically incorrect.

Disjunction and conjunction

English sentences of the form 'ϕ or ψ' are often uttered in a context where the available evidence does not enable the speaker to decide which of the sentences ϕ and ψ are true, but only tell him that at least one of these sentences *must* be true. Moreover, it would seem that sentences of the form 'ϕ or ψ' are sometimes true, and indeed true on the basis of the available evidence when uttered in such a context. Take for example the sentence

(a) *Either the red marble or the yellow marble is in box 2*

uttered by someone in information state 2. Clearly, there is nothing wrong with this statement, even though it is not yet settled which of these two marbles really is to be found in box 2.

If this observation is correct, it would seem that in most contexts the operator v cannot serve as the formal counterpart of 'or'. According to definition 2, a sentence of the form $(\phi \vee \psi)$ is not true on the basis of the available evidence unless it is clear which of the sentences ϕ and ψ is true on that basis - and, on most occasions, this is a bit too much to ask.

Fortunately, the present theory provides yet another possible analysis of disjunctive sentences: in place of a sentence of the form $(\phi \vee \psi)$ one can take a sentence of the form *must*$(\phi \vee \psi)$ as their formal translation. *must*$(\phi \vee \psi)$ is true on the basis of the evidence available at a given information state s iff this information state cannot possibly grow into an information state s' where both ϕ and ψ are false on the basis of the available evidence. This means that at least one of the sentences ϕ and ψ will eventually turn out to be true if one continues to accumulate information.

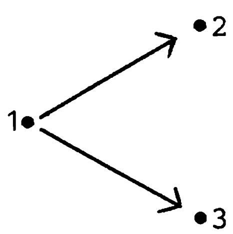
At this point the reader may wonder why I did not assign to sentences of the form $(\phi \vee \psi)$ the truth conditions which are now associated with sentences of the form *must* $(\phi \vee \psi)$. Wouldn't that have been a more elegant procedure?

The reason I did not proceed that way is this: sometimes disjunction *is* used in the manner formally captured by the truth and falsity conditions associated with the operator $\vee$. In fact, from a syntactical point of view, there are only a few cases - the case where 'or' occurs as the only connective of the relevant sentence being the most salient - in which the meaning of English disjunction does not seem to conform to the meaning of $\vee$. Yet I venture the hypothesis that even in these special cases the *literal* meaning of 'or' can be equated with the meaning of $\vee$, and that it is for *pragmatic* reasons that a sentence of the form 'ϕ or ψ' often has to be interpreted as 'it must be the case that ϕ or ψ'.

I will not defend this position here. Fred Landman will do so in his forthcoming dissertation. In the meantime the reader is invited to think of 'better' clauses for $\vee$. (This is not just a matter of changing the true clause for $\vee$ into

(*) $s \models \phi \vee \psi$ iff for no $s' \geq s$, $s \mathrel{=\!\!|} \phi$ or $s \mathrel{=\!\!|} \psi$

while leaving everything else as it is. An example will make this clear. Consider the information model given by the following picture.



$V_1(\phi)$ is undefined
$V_2(\phi) = 1$
$V_3(\phi) = 0$

Applying (*) and the clauses for *must* and $\sim$ given in definition 2, we get

$$1 \not\models \textit{must}\ \phi \vee \textit{must}\ \sim\phi$$

Applying the falsity clause for v and the clauses for *must* and ~ given in definition 2, we get

$$1 \dashv \textit{must}\ \phi \vee \textit{must}\ \phi$$

Obviously, something is wrong here.)

So much for disjunction. The truth and falsity conditions pertaining to conjunction need no further comment - except perhaps that much of what has been said about sentences of the form 'ϕ or ψ' applies equally well to sentences of the form 'not both ϕ and ψ'.

III.4. PROPOSITION. Suppose ~, ∧, v are the only operators occurring in ϕ. Then ϕ is stable.

PROOF. Induction on the complexity of ϕ. □

In the sequel I shall sometimes discriminate between the sentences which contain no operators other than ~, v and ∧, and the other ones by calling the former *descriptive* and the latter *nondescriptive*. All descriptive sentences are stable, most nondescriptive sentences are not. Intuitively, the difference between these two kinds of sentences amounts to this: by uttering a descriptive sentence a speaker only informs his audience of the evidence he already has. By uttering a nondescriptive sentence he also expresses his expectations about the outcome of further investigations.

III.5. PROPOSITION. Let $M = \langle S, \leq, V \rangle$ be a closed information model, and $s \in S$.

(i) For no sentence ϕ, both $s \models \phi$ and $s \dashv \phi$

(ii) If s is maximal, then for every sentence ϕ, either $s \models \phi$ or $s \dashv \phi$. In particular we find

$s \models \phi \rightarrow \psi$ iff $s \not\models \phi$ or $s \models \psi$

$s \models \textit{may}\ \phi$ iff $s \models \phi$

$s \models \textit{must}\ \phi$ iff $s \models \phi$

PROOF. Induction. □

What this proposition shows is that it does not make much sense to use the phrases 'if ... then', 'must', and 'may' in a context where the information is complete. In such a context 'if ... then' gets the meaning of the material implication while the meaning of both 'must' and 'may' boils down to that of the empty operator. However, in such a context there is no need to use non-descriptive sentences: the information is complete; so, what could possibly be the good of speculations on the outcome of *further* investigations?

Proposition 5 also enables us to clarify the relation between the relative notions 'true/false on the basis of the available evidence' and the absolute notions 'true' and 'false'. Indeed, the reader may have wondered whether these notions are related at all. Wouldn't it be better to say that definition 2 deals with the notions of verification and falsification rather than the notions of truth and falsity? After all, it is obvious that nothing is verified or falsified except on the basis of evidence. But it is far from obvious that this evidence, or rather the availability of it, could make a difference to the truth value of the sentence concerned. Truth and falsity depend only on the facts of the case and not on information one may have gathered. [5)]

The absolute notions of truth and falsity can be defined in terms of the relative notions as follows: a sentence is true/false iff it is true/false on the basis of the evidence that will be available when the data are complete. In formulas:

$$M \models \phi \text{ iff } s_o \models_M \phi$$
$$M \mathrel{=\!\!|} \phi \text{ iff } s_o \,{}_M\!\!\mathrel{=\!\!|}\, \phi$$

Here s_o is the special information state discussed near the end of section 1. The evidence available at s_o comprises exactly what is in fact the case. Hence, it is indeed the facts and nothing but the facts that determine whether a sentence is true or false in the absolute sense.

We saw, however, that there are many sentences for which the absolute notions of truth and falsity make little sense.

There is a lot to learn from what is in fact the case but not which sentences may be true or must be true or will be true if only ... There is no way to decide the question whether the red and the yellow marble *may* both be in box 1 by just opening the boxes. A question like that can only be judged in the light of what *may* be case: the possibilities left open by the facts as far as they are known.

Given the possibilities left open by the facts known in information state 2, the yellow and the red marble cannot both be in box 1. The sentence 'the yellow and the red marble may both be in box 1' is false on the basis of the evidence available in information state 2. Now, I have no objections against replacing this phrase by another one - 'falsified by the available evidence' or 'refutable in information state 2', whatever you like. The real issue is, I think, which notions are fundamental: the absolute notions of truth and falsity or the relative ones, whatever you call them. In this paper we are exploring the idea that the relative notions are fundamental. So far it has proven fairly fruitful: it enables us to draw the distinction between direct and indirect evidence and that between stable and unstable sentences - important distinctions it would seem, even in purely logical matters.

III.1.3. Data logic. Preliminaries

III.6. DEFINITION. Let ϕ be a sentence and Δ a set of sentences $\Delta \models \phi$ iff every closed information model $M = \langle S, \leqq, V \rangle$ is such that for every $s \in S$,

if $s \models_M \psi$ for every $\psi \in \Delta$, then $s \models_M \phi$. □

'$\Delta \models \phi$' abbreviates 'the argument Δ/ϕ is (data-)logically *valid*'. I shall write '$\models \phi$' instead of '$\emptyset \models \phi$', and '$\Delta, \psi_1, \ldots, \psi_n \models \phi$' instead of '$\Delta \cup \{\psi_1, \ldots, \psi_n\} \models \phi$'. Read '$\models \phi$' as '$\phi$ is logically valid'.

The logic generated by the above definition differs in some important respects from any in the literature. Most of the following observations have to do with these differences. For a more systematic account the reader is referred to chapter 3.

And, or and *not*

As far as descriptive sentences are concerned, the departure from classical logic is not too drastic. Many classical principles concerning $\wedge$, $\vee$ and $\sim$ are valid in the sense of definition 6 as well:

- If $\Delta \models \phi \wedge \psi$, then $\Delta \models \phi$ and $\Delta \models \psi$;
- If $\Delta \models \phi$ and $\Delta \models \psi$ then $\Delta \models \phi \wedge \psi$;
- If $\Delta \models \phi \vee \psi$, and $\Delta, \phi \models \chi$, and $\Delta, \psi \models \chi$, then $\Delta \models \chi$;
- If $\Delta \models \phi$ or $\Delta \models \psi$, then $\Delta \models \phi \vee \psi$;
- If $\Delta \models \sim\phi$ and $\Delta \models \phi$, then $\Delta \models \psi$;
- If $\Delta \models \sim\sim\phi$ then $\Delta \models \phi$.

Notice that this list is made up entirely of principles underlying the classical system of natural deduction for $\wedge$, $\vee$, and $\sim$. Actually, only one of these principles is missing. Within the present context the usual introduction rule for negation fails. It is not necessarily so that

- If $\Delta, \phi \models \sim\phi$ then $\Delta \models \sim\phi$.

The closest approximation available is this:

- If $\Delta, \phi \models \sim\phi$, and each $\psi \in \Delta$ is T-stable, then $\Delta \models$ *must* $\sim\phi$.

The matter can also be put as follows. Within the present context proofs by *Reductio ad Absurdum* are not always valid. It is not generally so that

- If $\Delta, \sim\phi \models \psi \wedge \sim\psi$ then $\Delta \models \phi$.

That is, if you can derive an absurdity from the assumption $\sim\phi$ this does not always qualify as *direct evidence* for ϕ. At best it gives you *indirect evidence* for ϕ, but even this only in those cases where the premises in Δ are all T-stable. So we get

- If $\Delta, \sim\phi \models \psi \wedge \sim\psi$ and each $\psi \in \Delta$ is T-stable, then $\Delta \models$ *must* ϕ.

III.7. PROPOSITION. Let $\wedge$, $\vee$, and $\sim$ be the only operators occurring in the sentences of Δ/ϕ. Suppose Δ/ϕ is classically valid. Then $\Delta \models$ *must* ϕ.

PROOF. Given propositions 4 and 5, there is little left to prove.

In other words, if by the standards of classical logic, the descriptive sentence ϕ must follow from the descriptive sentences in Δ, then at least *must* ϕ follows from Δ by the standards set here. (In this respect *must* behaves in data logic as double negation in intuitionistic logic.)

Example. Let ϕ be descriptive. Then $\not\models \phi \vee \sim\phi$. The Principle of Excluded Middle does not generally hold. (In fact there are no valid descriptive sentences at all.) That does not mean, however, that $\phi \vee \sim\phi$ can ever be false on the basis of the available evidence: *must*$(\phi \vee \sim\phi)$ is logically valid. Besides, we get a Principle of Excluded Muddle in return. No matter what the exact evidence is, the sentence

must $\phi \vee$ *must* $\sim\phi \vee ($*may* $\phi \wedge$ *may* $\sim\phi)$

is always true on the basis of it.

Must and *may*

In many respects *must* and *may* behave as ordinary modal operators. We get for example

must $\phi \models \sim$*may* $\sim\phi$

$\sim$*may* $\sim\phi \models$ *must* ϕ

which shows that *must* and *may* are related as the box $\Box$ and the diamond $\Diamond$ of any old modal system. We also have

must$(\phi \wedge \psi) \models$ *must* $\phi \wedge$ *must* ψ

must $\phi \wedge$ *must* $\psi \models$ *must*$(\phi \wedge \psi)$

$\models$ *must*$(\phi \vee \sim\phi)$

which makes it seem as if we are dealing with just another extension of $\mathbf{K}$. (See section II.2.3.) Furthermore, we have

must $\phi \not\models \phi$

must $\phi \models$ *may* ϕ

must $\phi \models$ *must must* ϕ

must must $\phi \models$ *must* ϕ

which might easily create the impression that *must* and *may* are the obligation and permission operators of some system of *deontic* logic.

But then we see

$\phi \models$ *may* ϕ

which not only gives the logic of *may* an *alethic* flavour, but also seems irreconcilable with the earlier observations.

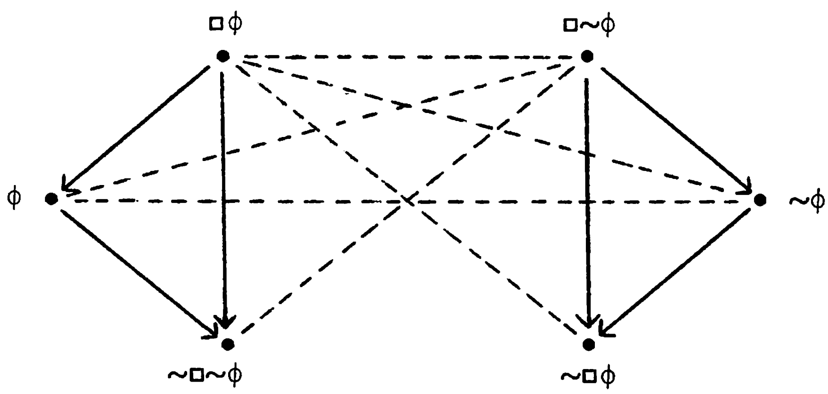
And finally, we note

If ϕ is T-stable, then $\phi \models$ *must* ϕ

If ϕ is F-stable, then *may* ϕ, $\sim\phi \models \chi$ for any χ

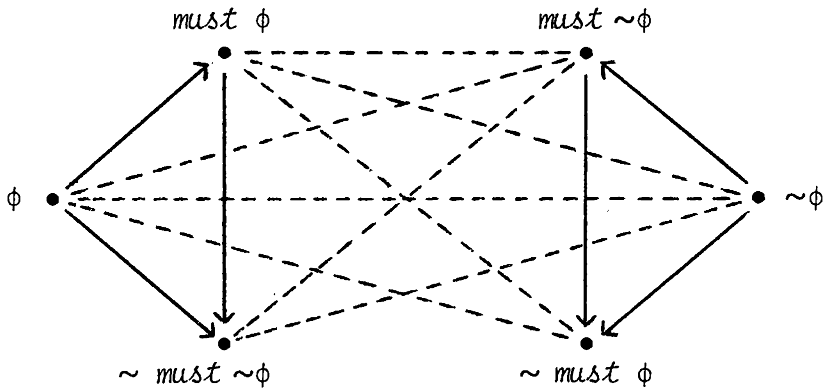
which distinguish this logic clearly from all earlier systems. (Cf. the discussion in section 1.)

The following two figures will help to clarify the situation



What is pictured here are the logical relations between the six formulas ϕ, $\sim\phi$, $\Box\phi$, $\Box\sim\phi$, $\sim\Box\phi$, $\sim\Box\sim\phi$ in any of the standard alethic modal logics. I have drawn an arrow from ψ to χ to indicate that $\psi \models \chi$. When ψ and χ are connected by a dotted line this means that $\psi, \chi \models \Theta$ for any Θ.

Now, if we draw a similar picture for *must* we get this, at least if ϕ is a *stable* formula.



Here, the assumption that ϕ is *stable* is essential. For example, if ϕ = *may* ψ, for any atomic ψ we find

may $\psi \not\models$ *must may* ψ

In fact, things are the other way around:

must may $\psi \models$ *may* ψ

This last example shows that the Principle of Substitution cannot be carried over from classical logic to data logic without modification. It may very well be that a given argument is valid - ϕ / *must* ϕ, for instance - while the result of substituting an *arbitrary* sentence - *may* ψ - for some atomic sentence (at all places where this atomic sentence occurs) is not valid any more - witness *may* ψ / *must may* ψ. In general, only substitution of a *stable* sentence for an atomic sentence will transform a valid argument into a valid one.

Also the Principle of Replacement needs to be treated with some care. Let us call the sentences ϕ and ψ *weakly equivalent* iff both $\phi \models \psi$ and $\psi \models \phi$, and *strongly equivalent* iff $\phi \models \psi$, $\psi \models \phi$, $\sim\phi \models \sim\psi$ and $\sim\psi \models \sim\phi$. This distinction is important. Consider, for example, the sentences $(\phi \wedge \sim\phi)$ and $(\psi \wedge \sim\psi)$, where ϕ and ψ are two distinct atomic sentences. $(\phi \wedge \sim\phi)$ and $(\psi \wedge \sim\psi)$ are weakly equivalent but not strongly equivalent. If the occurrence of $(\phi \wedge \sim\phi)$ in $\sim(\phi \wedge \sim\phi)$ is replaced by an occurrence of $(\psi \wedge \sim\psi)$ then the resulting sentence $\sim(\psi \wedge \sim\psi)$ is *not* weakly equivalent to the original $\sim(\phi \wedge \sim\phi)$. Hence, the Principle of Replacement fails for weak equivalents. Yet, it holds for strong equivalents: if ϕ and ψ are strongly equivalent then replacement of an occurrence of ϕ in a sentence χ by an occurrence of ψ will always yield a sentence χ' that is strongly equivalent to the orginal χ.

Examples

$\sim\sim\phi$ is strongly equivalent to ϕ
$\phi \vee \psi$ is strongly equivalent to $\sim(\sim\phi \wedge \sim\psi)$
$\phi \wedge \psi$ is strongly equivalent to $\sim(\sim\phi \vee \sim\psi)$
may ϕ is strongly equivalent to $\sim$ *must* $\sim\phi$
must ϕ is strongly equivalent to $\sim$ *may* $\sim\phi$
may ϕ is strongly equivalent to $\sim(\phi \rightarrow \sim\phi)$
must ϕ is strongly equivalent to $\sim\phi \rightarrow \phi$
$\phi \rightarrow \psi$ is strongly equivalent to *must*$(\sim\phi \vee \psi)$
$\phi \rightarrow \psi$ is strongly equivalent to $\sim$*may*$(\phi \wedge \sim\psi)$

So we see that in principle it is possible to give a more economic presentation of the present system by choosing $\sim$, one of $\wedge$ and $\vee$, and one of *may*, *must*, and $\rightarrow$ as primitive operators and defining the other ones in terms of these.

Let us return to the logic of *may* and *must*. As usual [6] we define a *modality* as any unbroken sequence of zero or more monadic operators ($\sim$, *must*, *may*). Two modalities X and Y are *equivalent* iff the result of replacing X by Y (or Y by X) in any sentence is always strongly equivalent to the original sentence. Proposition 8 says how many equivalence classes there are.

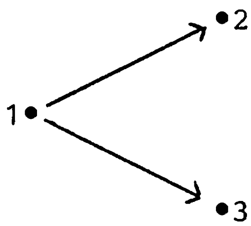
III.8. PROPOSITION. Let X be any modality and ϕ be any sentence.

(i) $X\phi$ is strongly equivalent to one or other of the following *ten* sentences:
ϕ, *must* ϕ, *may* ϕ, *must may* ϕ, *may must* ϕ,
$\sim\phi$, $\sim$*must* ϕ, $\sim$*may* ϕ, $\sim$*must may* ϕ, $\sim$*may must* ϕ

(ii) If ϕ is *stable*, $X\phi$ is strongly equivalent to one or other of the following *six* sentences:

ϕ, *must* ϕ, *may* ϕ
$\sim\phi$, $\sim$*must* ϕ, $\sim$*may* ϕ

PROOF. Below a model is pictured showing that *must may* ϕ is not always strongly equivalent to *must* ϕ.



$V_1(\psi)$ is undefined
$V_2(\psi) = 1$
$V_3(\psi) = 0$

Note that $1 \models$ *must may* (*must* ψ $\vee$ *must* $\sim\psi$), and $1 =\!|$ *must* (*must* ψ $\vee$ *must* $\sim\psi$).
However, if ϕ is F-stable, *must may* ϕ is strongly equivalent to *must* ϕ.

The picture also shows that *may must* ϕ is not always strongly equivalent to *may* ϕ, witness the fact that $1 \models$ *may*(*may* ψ $\wedge$ *may* $\sim\psi$), and $1 =\!|$ *may must* (*may* ψ $\wedge$ *may* $\sim\psi$).
However, if ϕ is T-stable, *may must* ϕ is strongly equivalent to *may* ϕ.

Turning to *may must may* ϕ, we find that a sentence of this form is always strongly equivalent to *may must* ϕ, no matter the stability of ϕ. From this it follows immediately that *must may must* ϕ is always strongly equivalent to *must may* ϕ.

The remainder of the proof is left to the reader. □

In section 4.2, I will propose a somewhat subtler analysis of *may* and *must* - one which results in (ii) holding even for unstable ϕ.

16

III.9. PROPOSITION. Suppose that $\rightarrow$ is the only operator occurring in ϕ. Then,

$$\models \phi \text{ iff } \phi \text{ is classically valid}$$

PROOF. The proof from right to left is trivial. The converse follows straightforwardly from the following lemma.
Let $M = \langle S, \leq, V\rangle$ be any closed information model, and let ϕ be any sentence in which no operator occurs other than $\rightarrow$. Suppose there is some $s \in S$ such that $s \dashv \phi$. Then there is some *maximal* $s' \in S$ such that $s' \geq s$ and $s' \dashv \phi$. □

This suggests that the logic attributed to indicative conditionals by the theory presented here is rather strong. The scope of proposition 9 is, however, very limited. For one thing, one cannot even conclude from it that if $\rightarrow$ is the only operator occurring in ϕ and in the sentences of Δ, it holds that

$$\Delta \models \phi \text{ iff } \Delta/\phi \text{ is classically valid}$$

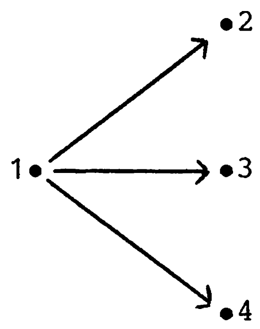
Example. Let ϕ and ψ be atomic. Then the argument $\phi, \phi \rightarrow \psi / \psi$ being an instance of *Modus Ponens*, is classically valid. According to definition 2, however, $\phi \rightarrow \psi$ is true on the basis of the data just in case these data augmented with *direct* evidence for ϕ amount to *indirect* evidence for ψ. So we get $\phi, \phi \rightarrow \psi \models \textit{must}\ \psi$ rather than $\phi, \phi \rightarrow \psi \models \psi$.

III.10. PROPOSITION. Let $\rightarrow$ be the only operator occurring in ϕ and in the sentences of Δ.
If Δ/ϕ is classically valid, then $\Delta \models \textit{must}\ \phi$.

PROOF. The proposition is another corollary of the lemma mentioned in the proof of proposition 9. □

Both proposition 9 and proposition 10 remain true if we permit the conjunction operator to occur in ϕ and in the sentences of Δ. For other connectors things go wrong.

Example. Let ϕ, ψ, χ be atomic sentences. The argument $(\phi \wedge \psi) \rightarrow \chi \ /\ (\phi \rightarrow \chi) \vee (\psi \rightarrow \chi)$ is classically valid. Within data logic, however, it isn't. Nor is the argument $(\phi \wedge \psi) \rightarrow \chi \ /\ \textit{must}((\phi \rightarrow \chi) \vee (\psi \rightarrow \chi))$, witness the model pictured below.



$V_2(\phi) = V_3(\phi) = 1;\ V_4(\phi) = 0$
$V_3(\psi) = V_4(\psi) = 1;\ V_3(\psi) = 0$
$V_3(\chi) = V_4(\chi) = 0;\ V_2(\chi) = 1$

Clearly $1 \models (\phi \wedge \psi) \rightarrow \chi$; $1 =\!| (\phi \rightarrow \chi) \vee (\psi \rightarrow \chi)$; and $1 =\!| \mathit{must}((\phi \rightarrow \chi) \vee (\psi \rightarrow \chi))$.
The model also explains what is wrong with the following informal example, taken from Adams (1975):

If switches A and B are thrown, the motor will start

∴ *(It must be the case that) either if switch A is thrown, the motor will start, or, if switch B is thrown the motor will start*

Not only the principle of Modus Ponens, but also the other mainstay of classical logic, the principle of Free Deduction (see section II.2.1), cannot be carried over to data logic without modification. It is not generally so that

If $\Delta, \phi \models \psi$ then $\Delta \models \phi \rightarrow \psi$.

If you want to prove $\phi \rightarrow \psi$ from the premises Δ, you cannot just add ϕ to the premises and try to prove ψ from Δ together with ϕ. By making an assumption - suppose that ϕ will turn out true - you may rule out some of the possibilities left open by the premises - *may* $\sim\phi$, for example. In other words, the assumption ϕ can interfere with the T-unstable sentences in Δ. Therefore, if you want to prove ψ from the assumption ϕ, you may only use the T-stable premises. So, we get

If $\Delta, \phi \models \psi$ and each $\chi \in \Delta$ is T-stable, then $\Delta \models \phi \rightarrow \psi$.

(Actually, we can get something a bit stronger, for the premises Δ together with the assumption ϕ have to supply only *indirect* evidence for ψ:

If $\Delta, \phi, \sim\psi \models \psi$, and each $\chi \in \Delta$ is T-stable,
then $\Delta \models \phi \rightarrow \psi$.)

By now it will be clear that the logic attributed to indicative conditionals by the theory presented here cannot easily be fitted into the spectrum formed by the theories proposed so far. In certain respects it is very strong, as strong as classical logic: recall that $\phi \rightarrow \psi$ follows from ψ, and also from $\sim\phi$, at least if ϕ and ψ are stable. In other respects the behaviour of $\rightarrow$ matches with the behaviour of the strict implications occurring in the Lewis Systems (I mean C.I. Lewis here): as we saw $\sim(\phi \rightarrow \psi)$ is (strongly) equivalent with *may*$(\phi \wedge \sim\psi)$. This is exactly what we would get if $\rightarrow$ were the implication and *may* the possibility operator of another extension of S.0.5. In yet other respects data logic is weaker than the weakest logic in the literature. One more example: the principle of Modus Tollens, which holds both in classical and in intuitionistic logic, and also in the systems of strict and variably strict implication, and even in such a weak system as the system of Relevance Logic, fails. One cannot in general conclude $\sim\phi$ from $\phi \rightarrow \psi$ and $\sim\psi$. The closest approximation available is this: if ψ is F-stable then *must* $\sim\phi$ follows from $\phi \rightarrow \psi$ and $\sim\psi$. If ψ is not F-stable even this weakened version of Modus Tolles does not hold. Consider for example the premises $\phi \rightarrow (\psi \rightarrow \chi)$ and $\sim(\psi \rightarrow \chi)$, where ϕ, ψ, and χ are three distinct descriptive sentences. Neither $\sim\phi$ nor *must* $\sim\phi$ follow from these sentences, we only have that *may* $\sim\phi$ is true on the basis of the available evidence if $\phi \rightarrow (\psi \rightarrow \chi)$ and $\sim(\psi \rightarrow \chi)$ are. (You will have recognized the case of the marbles. See also sections I.1.1 and I.2.1.2.)

III.2. PRAGMATIC CORRECTNESS AND INCORRECTNESS

In certain respects data logic is weaker than the weakest logic in the literature. In other respects it is at least as strong as any of the others.
I like to think that the arguments which on my account are logically invalid cannot easily be explained away as 'just' pragmatically unsound by those who think they are valid. Notoriously difficult (for those who believe that indicative conditionals behave like material implications) are for example the schemes $\sim(\phi \rightarrow \psi)\ /\ \phi$ and $(\phi \wedge \psi) \rightarrow \chi\ /\ (\phi \rightarrow \chi) \vee (\psi \rightarrow \chi)$.
So far no satisfactory pragmatic explanation has been offered for the fact that many instances of these inference patterns seem anomalous.

On the other hand, those who think that my theory is too strong, that I have dubbed too many of the wrong arguments valid, can produce a lot of intuitive counterexamples to make their point. Here I am the one who has to produce the good reasons for saying that these are 'just' pragmatically unsound instances of valid argument forms. I shall turn to this now.

III.2.1. Gricean constraints

Recall the Principle of Excluded Muddle:
must $\phi \vee$ *must* $\sim\phi \vee$ (*may* $\phi \wedge$ *may* $\sim\phi$) is a datalogical law. This means that the possible contexts in which a conditional with antecedent ϕ and consequent ψ can be uttered all fall into the following nine categories:

1. *must* ϕ *must* ψ	2. *must* ϕ *may* ψ *may* $\sim\psi$	3. *must* ϕ *must* $\sim\psi$
4. *may* ϕ *may* $\sim\phi$ *must* ψ	5. *may* ϕ *may* $\sim\phi$ *may* ψ *may* $\sim\psi$	6. *may* ϕ *may* $\sim\phi$ *must* $\sim\psi$
7. *must* $\sim\phi$ *must* ψ	8. *must* $\sim\phi$ *may* ψ *may* $\sim\psi$	9. *must* $\sim\phi$ *must* $\sim\psi$

(Read this as follows: in category 1 *must* ϕ is true on the basis of the evidence available to the speaker and *must* ψ too. Etc.)

Claim: Assume that ϕ and ψ are descriptive sentences. Then the only contexts in which a speaker can assert $\phi \rightarrow \psi$ without violating any conversational maxim are the ones in category 5. In other words, an indicative conditional statement with a descriptive antecedent and consequent will normally implicate that neither the truth nor the falsity of its antecedent or consequent are definitely established.

The claim itself is not new. Already in Strawson (1952: 88) we find the remark that 'the hypothetical statement carries the implication either of uncertainty about, or disbelief in, the fulfillment of both the antecedent and consequent'. (See also Stalnaker (1976) and Gazdar (1979).) What is new is the straightforward proof of it. Consider first the contexts fitting into category 2, 3 or 6. In such contexts the sentence $\phi \rightarrow \psi$ is false on the basis of the evidence available to the speaker. (It is left to the reader to check this with the help of definition 2.) So if you say $\phi \rightarrow \psi$ in one of these contexts you are saying something for which you lack adequate evidence, which according to the maxim of quality (the one formulated in section 1.2) you are not supposed to do. Secondly, anyone who knows, or at least could have known, that ϕ cannot be true, and who therefore falls within one of the categories 7, 8 or 9 could according to definition 2 truthfully assert that $\phi \rightarrow \psi$. But anyone who did so would be sinning against the maxims of quantity and manner: by definition 2. *must* $\sim\phi$ is stronger and therefore more informative than $\phi \rightarrow \psi$. Apart from that it is also less wordy. So, it would be a lot more helpful to say *must* $\sim\phi$. The only remaining categories are 1 and 4, in both of which the speaker knows that ψ must be the case. Again: *must* ψ is both stronger and less wordy than $\phi \rightarrow \psi$. So, if you say $\phi \rightarrow \psi$ in such a context, you are not telling us all you know and that in too many words.

So, indicative conditionals are typically uttered in contexts fitting in category 5, the center of the table. This is of course not to say that any conditional statement will automatically be correct when uttered in such a context. For one thing, in such a context the sentence $\phi \rightarrow \psi$ cannot be true on the basis of the available evidence unless the antecedent ϕ is somehow 'relevant' to the consequent ψ. Let ψ be any descriptive sentence - take 'the red marble is in box 1'. Suppose you do not know whether ψ - maybe the red marble is in box 1, maybe not. Likewise, let ϕ be any descriptive sentence - 'it is raining in Ipanema'. Again, you do not know whether ϕ - maybe it is raining in Ipanema, maybe not. Now consider $\phi \rightarrow \psi$ - 'if it is raining in Ipanema, the red marble is in box 1'. Clearly, there must be some non-coincidental connection between ϕ and ψ if it is really to be so that no additional evidence can establish the truth of ϕ, without establishing that ψ must be true - how on earth could the weather condition in Ipanema have anything to do with the position of the marbles?

In section 1.2 we noted that our semantics in itself does not guarantee that a conditional is true on the basis of the available evidence if its antecedent is relevant to its consequent. We can now see why this does not matter too much. Pragmatic constraints ensure that an indicative conditional will normally be asserted only in circumstances where this requirement is fulfilled. Those contexts in which definition 2 makes a conditional true without the antecedent being relevant to the consequent are contexts in which so much is known about the truth and falsity of either of these that it cannot be asserted without violating some conversational maxim.

III.2.2. Odd conditionals

Should conditionals never be uttered in other circumstances than the ones fitting in category 5, just because this violates one or the other conversational maxim? Of course not. There are plenty of good occasions for doing just this, only it must be clear that a maxim has been overruled and why.

Contexts fitting into 2, 3 and 6 are not among these occasions. There the conditional is false on the basis of the evidence available to the speaker and as we noticed in section 1.2 any violation of the maxim of quality is incompatible with the overall Cooperative Principle. [7)]

But the literature is full of if's and then's with the most eccentric things in between and all those I know fit quite neatly in that part of the table formed by the categories 1, 4, 8 and 9. In fact, this categorization is of great help when we want to classify the figures of speech beginning with *if*.

All of the examples which go

(a) *If ..., I'll eat my hat*

belong to category_9: the speaker is clearly not intending to eat her hat and the hearer is expected to complete the (weakened version of) Modus Tollens for herself, which gives

(b) *It cannot be the case that ...*

Why say (a) rather than (b)? Surely in order to make the claim that the antecedent is as definitely false as the applied Modus Tollens is valid. The same rhetoric occurs in constructions like

(c) *If ..., I am a Dutchman*
(d) *If ..., I am the Empress of China*
(e) *I'll be hanged, if ...*

which all implicate the falsity of their antecedents (unless of course the speaker could be a Dutchman, or the Empress of China, or sentenced to death).

There are also plenty of examples of which the antecedent is trivially true and the hearer is supposed to apply (the weakened version of) Modus Ponens:

(f) *She is on the wrong side of thirty, if she is a day*
(g) *If there is one thing I cannot stand, it is getting caught in the rushhour traffic*

It will be clear that these examples belong to category_1.

Category_4 is the most diverse. In addition to examples where 'if' is used for purely rhetorical reasons, like

(h) *This is the best book of the month, if not of the year* it also contains examples where *if* serves as an opting out device.

(i) *There is coffee in the pot, if you want some*
(j) *If there is anything you need, my name is Marcia*
(k) *I paid back that fiver, if you remember*
(l) *If I may interrupt you, you're wanted on the telephone*

Let us first discuss (h). The speaker supposes that the hearer is well aware of the trivial truth that this book will certainly be the best of the month if it *is* the best of the year. In formulas, the hearer is supposed to know that $\phi \rightarrow \psi$. From this together with what the speaker tells, $\sim\phi \rightarrow \psi$, the hearer could (by data logical means) conclude *must* ψ: this must be the best book of the month. Just as in the above examples the speaker intends the hearer to draw this conclusion.

Example (i) works differently. The hearer knows that the speaker is not in the position to know whether the hearer wants some coffee or not. From this the hearer can infer that the conditional is asserted in one of the categories 4, 5 or 6. It cannot be category 6, for then the statement would be false on the basis of the information available to the speaker. For the same reason it cannot be category 5 (unless the speaker happens to be a genie who could just make coffee in the pot on command - but let us assume that the hearer knows this is not the case). So the only possibility left is category 4: there must be coffee in the pot.

To what good purpose - if any - does the speaker prefer the *if*-form to the statement that *there is coffee in the pot*? I think that the speaker in simply asserting the consequent would run the risk of defying the maxim of relevance, by saying something which does not interest the hearer at all. With the antecedent the speaker indicates that he is well aware of this: it provides a condition under which the consequent will be interesting. The examples (k) and (l) show that it is not always the maxim of relevance that is involved. In (k) the speaker indicates with the antecedent that he is opting out of the maxim of quantity [8]; to account for (l) we must appeal to a maxim of politeness [9][10].

Also in category 8 one can breach the conversational maxims to good effect:

(m) *If it does not rain tomorrow, then it is going to pour*

(given as a summary of a dismal weather forecast)

(n) *If I don't beat him, I'll thrash him*

(a boxer boasting before his fight). Both (m) and (n) convey that their antecedent will turn out false, but they leave their consequent undecided. The reader will be able to work out these implicatures himself. (m) and (n) both mirror example (h).

I have not been able to find any good (idiomatic) conditionals fitting in category 7. Nor can I offer a satisfactory explanation why there aren't any. A rather unsatisfactory explanation runs like this: saying $\phi \rightarrow \psi$ and conveying by this both the more informative *must* $\sim\phi$ and the more informative *must* ψ involves violating the maxim of quantity not once but twice. It could be asking too much of a hearer to expect him to work this out.

The examples discussed above must look odd if not perplexing to those who hold the view that a conditional statement cannot be true unless the antecedent and the consequent are in some sense 'causally' connected. How could any causal chain ever bridge the gap between the antecedent *she is a day* and the consequent *she is on the wrong side of thirty* of (f); or that between the antecedent *there is anything you need* and the consequent *my name is Marcia* of (j)? Given that how the dots are filled in is irrelevant to the truth of *if ..., I'll eat my hat* as long as they are filled in with something which is false, what could such a sentence express if not a simple truth functional connection between the antecedent and the consequent?

One might reply that these examples only show that the *if* of natural language is ambiguous: usually it expresses a causal connection, but in some exceptional cases it does not. I do not think that this is the only way to see it. One of the advantages of the data semantic approach is that we can uphold the idea of an unambiguous *if*. The *if* that enables a speaker (in information state 1) to formulate the general constraint that the blue marble is in box 2 if the other two are in box 1 is the very same *if* that enables him (in information state 2) to say that the blue marble is in box 1 if it is anywhere at all.

Nonetheless, there are some distinctions which can usefully be drawn. I would distinguish conditionals that express a *law* from those expressing a *contingent* truth, and *normal* conditionals from *degenerate* ones.

Let $\phi \rightarrow \psi$ be an indicative conditional with a descriptive antecedent and consequent. Assume $\phi \rightarrow \psi$ is true on the basis of the evidence available at the information state s. Now, it might be that $\phi \rightarrow \psi$ really owes its truth to the evidence available at s: there are information states $s' \leqq s$, i.e. information states in which less data are available, such that $\phi \rightarrow \psi$ is false on the basis of the evidence available at s'. In this case $\phi \rightarrow \psi$ is a contingent truth. In the other case $\phi \rightarrow \psi$ expresses a law: the data available in s hardly matter; $\phi \rightarrow \psi$ would have been true even if no specific information had been available; $\phi \rightarrow \psi$ owes its truth to the structure of reality rather than to the particular facts which constitute it.[11)]

Both laws and contingent truths can be degenerate. They are so when nobody could truthfully assert them in a context fitting into category 5. More precisely: suppose $\phi \rightarrow \psi$ is true on the basis of the evidence available at s. $\phi \rightarrow \psi$ is degenerate in s just in case for any $s' \leqq s$, if $\phi \rightarrow \psi$ is true in s' then this is so because the evidence available in s' excludes either ϕ's truth or ψ's falsity (i.e, *must* $\sim\phi$ or *must* ψ is true on the basis of the evidence available at s').

Otherwise, and this is where one might say that there is something more than a truth-functional connection between antecedent and consequent, $\phi \to \psi$ is normal.

In information state 2 the sentence *If the yellow marble is in box 1, it isn't in box 2* is a normal law, and the sentence *If the yellow marble is in box 1, the red one is in box 2* a normal contingency; *If the yellow marble is both in box 1 and box 2, then so is the blue one* is a degenerate law, and *The blue marble is in box 1 if it is anywhere at all* a degenerate contingency.[12] Actually, many of the examples discussed in this section will be degenerate contingencies in any information state in which they are true on the basis of the available evidence.

III.2.3. A test for pragmatic correctness

Consider the following well-known example.

(a) *If there is sugar in the coffee then it will taste good*

∴ *If there is sugar in the coffee and diesel-oil as well then it will taste good*

This argument sounds suspicious. In fact, it is often claimed that it is quite possible to accept the premise while rejecting the conclusion. So it would seem that (a) provides a clear cut counterexample to the principle of *Strenghtening the Antecedent*. But is it really so clear cut? Compare (a) with (b).

(b) *Maybe there is diesel-oil in the coffee*

If there is sugar in the coffee, then it will taste good

∴ *If there is sugar in the coffee and diesel-oil as well, then it will taste good*

Would those who accept the premise of (a) while rejecting its conclusion also be prepared to accept the premises of (b)? I do not think so. Sugared coffee which may contain diesel-oil as well does not in general taste good. Yet the difference between (a) and (b) is very small: it is a conversational implicature of the conclusion of (a) that the coffee may well contain diesel-oil (along with other things you usually take in your coffee). All we did to get (b) was to add this implicature as another premise to the original argument. The addition of this implicature is enough to destroy the credibility of what is now the second premise: if there is sugar in the coffee, then it will taste good.

The same trick can be applied to the Smith/Jones example of section I.2.2.

(c) *If Jones wins the election, Smith will retire to private life*
If Smith dies before the election, Jones will win it

∴ *If Smith dies before the election, he will retire to private life*

Here the conclusion implicates that Smith may die before the election.

(d) *Maybe Smith dies before the election*
If Jones wins the election, Smith will retire to private life
If Smith dies before the election, Jones will win it

∴ *If Smith dies before the election, he will retire to private life*

As in the first example, the premises of the original argument (c), in particular the first premise, are no longer plausible once you are confronted with the implicatures of the argument's conclusion. Once you reckon with the possibility of Smith's sudden death, you are not likely to accept that he will retire if Jones should win the election. It may very well be that Jones wins the election and that Smith does not retire because he died.

From a data-logical point of view (a) and (c) are perfectly in order: if the premises are true on the basis of the available evidence, then so is the conclusion. But unfortunately the premises of these arguments cannot be true on the basis of the available evidence if one takes the pragmatic implicatures of the conclusion into account. Tł t is why they lack any cogency. I would suggest generalizing these examples as a test for the pragmatic correctness of an argument: *any argument of which the premises cannot hold if one takes the implicatures of the conclusion into account is pragmatically unsound.* This seems to be a reliable test for the following reason. The purpose of an argument is to convince others of its conclusion. You want to persuade someone to accept something he would perhaps rather not accept by showing that it logically follows from something he is willing to accept. In a way the conclusion comes first, *together with all its implicatures*, and the premises are brought in later when it appears that the conclusion is not taken for granted. But then, of course, it will not help if you bring in premises that are incompatible with the implicatures of the conclusion - unless you also say that in asserting the conclusion you have violated the conversational maxims.

There is one more argument which I have claimed to be pragmatically incorrect rather than logically invalid, and for which the above test yields the right result[13]; the marble example of section 1.2:

(e) *The blue marble is in box 1*

―――――――――――――――――――――――――――――

If the blue marble is in box 2, it is in box 1

The conclusion implicates that *the blue marble may not be in box 1*. If we add this to the premise *the blue marble is in box 1* we get a datasemantic contradiction. Hence (e) is not pragmatically correct. But the following argument, which is of the same form as (e), is without fault.

(f) *The blue marble is in box 1*

The blue marble is in box 1, if it is anywhere at all

Here the conclusion does *not* implicate that the blue marble may not be in box 1. It belongs to category 1 rather than to category 5.

Note that no instance of the argument form ψ / $\phi \rightarrow \psi$ will pass our test unless its conclusion is a *degenerate* conditional like in (f). In this respect the argument forms ψ / $\phi \rightarrow \psi$ and $\sim\phi$ / $\phi \rightarrow \psi$ differ from argument forms like the Hypothetical Syllogism $\psi \rightarrow \chi$, $\phi \rightarrow \psi$ / $\phi \rightarrow \chi$ and the principle of Strengthening the Antecedent $\phi \rightarrow \psi$ / $(\phi \wedge \chi) \rightarrow \psi$ which have many pragmatically correct instances with *normal* conclusions. Thus, in a way these so-called paradoxes of implication are paradoxes indeed, but only pragmatically so.

III.3. DATA LOGIC

III.3.1. Deduction principles

Data logic is not as messy as it may have appeared at first sight, when we compared it with the more established logics in section 1.3. As the more systematic investigations in this chapter will show, it turns out that it reduces to a fairly simple stock of deductive principles.

In the official parts of this chapter - the definitions, the lemmas, the theorems and their proofs - we will for reasons of economy assume that the object language has as its primitive operators only $\vee$, $\sim$ and *must*. Sentences of the form $(\phi \wedge \psi)$, *may* ϕ, and $(\phi \rightarrow \psi)$ will be treated as metalinguistic abbreviations of $\sim(\sim\phi \vee \sim\psi)$, $\sim$*must*$\sim\phi$, and *must* $(\sim\phi \vee \psi)$ respectively.

III.11. DEFINITION. Let Δ be any set of sentences. Then $\phi \in \Delta^t$ iff (i) ϕ is atomic; or (ii) $\phi = \sim\psi$ for some atomic ψ; or (iii) $\phi =$ *must* ψ for some ψ. □

The main reason why the sentences of Δ^t are of special interest is that they are all T-stable.

III.12. DEFINITION. Let $\mathfrak{D}$ be the smallest set of arguments for which the following holds.

- Id : $\phi/\phi \in \mathfrak{D}$ for every sentence ϕ
- Aug : if $\Delta/\phi \in \mathfrak{D}$, then $\Gamma/\phi \in \mathfrak{D}$ for every $\Gamma \supset \Delta$
- Cut : if $\Delta,\Gamma/\phi \in \mathfrak{D}$ and $\Delta/\psi \in \mathfrak{D}$ for every $\psi \in \Gamma$, then $\Delta/\phi \in \mathfrak{D}$

- ET v : if $\Delta/\phi \vee \psi \in \mathfrak{D}$, and $\Delta, \phi/\chi \in \mathfrak{D}$ and $\Delta, \psi/\chi \in \mathfrak{D}$, then $\Delta/\chi \in \mathfrak{D}$
- IT v : if $\Delta/\phi \in \mathfrak{D}$ or $\Delta/\psi \in \mathfrak{D}$, then $\Delta/\phi \vee \psi \in \mathfrak{D}$
- EF v : if $\Delta/\sim(\phi \vee \psi) \in \mathfrak{D}$, then $\Delta/\sim\phi \in \mathfrak{D}$ and $\Delta/\sim\psi \in \mathfrak{D}$
- IF v : if $\Delta/\sim\phi \in \mathfrak{D}$ and $\Delta/\sim\psi \in \mathfrak{D}$ then $\Delta/\sim(\phi \vee \psi) \in \mathfrak{D}$

- ET *must*: if $\Delta/\textit{must}\ \phi \in \mathfrak{D}$ and $\Delta/\sim\phi \in \mathfrak{D}$, then $\Delta/\psi \in \mathfrak{D}$
- IT *must*: if $\Delta^t, \sim\phi/\phi \in \mathfrak{D}$ then $\Delta/\textit{must}\ \phi \in \mathfrak{D}$
- EF *must*: if $\Delta/\sim\textit{must}\ \phi \in \mathfrak{D}$ and $\Delta/\textit{must}\ \phi \in \mathfrak{D}$, then $\Delta/\psi \in \mathfrak{D}$
- IF *must*: $\emptyset/(\textit{must}\ \phi \vee \sim \textit{must}\ \phi) \in \mathfrak{D}$

- EF~ : if $\Delta/\sim\sim\phi \in \mathfrak{D}$, then $\Delta/\phi \in \mathfrak{D}$
- IF~ : if $\Delta/\phi \in \mathfrak{D}$, then $\Delta/\sim\sim\phi \in \mathfrak{D}$

- M : if ϕ is atomic and $\Delta/\textit{must}\ \phi \in \mathfrak{D}$ then $\Delta/\textit{may}\ \phi \in \mathfrak{D}$ □

I have baptized this set of arguments $\mathfrak{D}$, because 'd' is the first letter in 'data logic'. We are, however, not entitled to think of $\mathfrak{D}$ as data logic until we have proved that $\mathfrak{D}$ really coincides with the set of all arguments that are valid in the sense of definition 6.

To facilitate comparison, I have specified $\mathfrak{D}$ in exactly the same way as I specified the logics figuring in part II. (See section II.1.1, in particular definition II.6 and the concluding remarks. Note in passing that $\mathfrak{D}$ is compact in the sense of definition II.13.)

For each of the operators *must* and v there are four deduction principles: an ET-rule, an IT-rule, an EF-rule, and an IF-rule. An ET-rule says how to *exploit* a sentence, once you know this sentence to be *true* on the basis of the information in Δ. For example, ETv says that if the disjunction $(\phi \vee \psi)$ follows from Δ, and if you want to show that χ follows from Δ, then it suffices to show that χ follows from Δ together with ϕ, and from Δ together with ψ. An IT-rule indicates how to *introduce* a *true* sentence. Thus, IT *must* says that if you want to show that *must* ϕ follows from Δ, it suffices to show that ϕ follows from Δ^t and the additional assumption $\sim\phi$. (If wanted, this rule can be liberalized a bit. There is nothing wrong with using other T-stable premises than the ones belonging to Δ^t in this deduction of ϕ. There are, however, theoretical reasons why the given rule is preferable: the more rigorous the rules, the stronger the completeness result will be.)

An EF-rule says how to *exploit* a *false* sentence in a deduction. EF$\vee$, for instance, says that once you know that $\phi \vee \psi$ is false on the basis of the information in Δ, you may conclude that both ϕ and ψ are false.

Finally, an IF-rule says how to proceed if you want to show that a sentence is false on the basis of the information in Δ. IF *must* does so by saying that for sentences of the form *must* ϕ the principle of Excluded Third holds. So if you can somehow show that *must* ϕ is not true on the basis of the information in Δ, you can be sure that $\sim$*must* ϕ is.

For the negation operator there is neither an IT-rule nor an ET-rule. Or rather, these are implicit in the IF-rules and the EF-rules for the other operators.

The M-rule falls outside the scope of the E- and I-rules. Its validity does not so much depend on the truth and falsity conditions of sentences of the form *must* ϕ, as on the structural constraints we imposed on the information models. (If wanted, M can be strengthened to a rule which says that for *any* ϕ, if Δ/*must* $\phi \in \mathfrak{D}$ then Δ/*may* $\phi \in \mathfrak{D}$.)
If we had started out with $\rightarrow$, $\vee$, and $\sim$ as primitive operators we would have included the following rules for $\rightarrow$.

- ET$\rightarrow$: if $\Delta/\phi \rightarrow \psi \in \mathfrak{D}$, $\Delta/\phi \in \mathfrak{D}$ and $\Delta/\sim\psi \in \mathfrak{D}$ then $\Delta/\chi \in \mathfrak{D}$;
- IT$\rightarrow$: if $\Delta^t,\phi,\sim\psi/\psi \in \mathfrak{D}$ then $\Delta/\phi \rightarrow \psi \in \mathfrak{D}$;
- EF$\rightarrow$: if $\Delta/\phi \rightarrow \psi \in \mathfrak{D}$ and $\Delta/\sim(\phi \rightarrow \psi) \in \mathfrak{D}$ then $\Delta/\chi \in \mathfrak{D}$;
- IF$\rightarrow$: $\emptyset/(\phi \rightarrow \psi) \vee \sim(\phi \rightarrow \psi) \in \mathfrak{D}$;
- M : if ϕ is atomic and $\Delta^t,\sim\phi/\phi \in \mathfrak{D}$, then $\Delta/\sim(\phi \rightarrow \sim\phi) \in \mathfrak{D}$.

Of course, if we had proceeded this way we should have defined Δ^t in such a way that it would contain all implications instead of all *must*-sentences.

III.13. PROPOSITION. Let ϕ be atomic.
if $\Delta/\phi \in \mathfrak{D}$ and $\Delta/\sim\phi \in \mathfrak{D}$, then $\Delta/\psi \in \mathfrak{D}$ for any ψ

PROOF. As in part II we use a Gentzen-style shorthand.

$$\dfrac{\dfrac{\dfrac{\Delta/\phi \qquad \dfrac{\dfrac{\phi/\phi}{\Delta^{t},\sim\phi,\phi/\phi}\,\text{Aug}}{\Delta,\phi/\mathit{must}\ \phi}\,\text{IT } \mathit{must}}{\Delta/\ \mathit{must}\ \phi}\,\text{Cut}}{\Delta/\sim\mathit{must}\sim\phi}\,\text{M} \qquad \dfrac{\dfrac{\dfrac{\sim\phi/\sim\phi}{\Delta^{t},\sim\sim\phi,\sim\phi/\sim\phi}\,\text{Aug}}{\Delta,\sim\phi/\ \mathit{must}\sim\phi}\,\text{IT } \mathit{must} \qquad \Delta/\sim\phi}{\Delta/\ \mathit{must}\sim\phi}}{\Delta/\psi}\,\text{EF } \mathit{must}$$

□

In fact, proposition 13 holds for any sentence ϕ. But since we do not need this result to prove the completeness of $\mathfrak{D}$, it is preferable to wait until a semantic proof can be given.

III.14. PROPOSITION. If $\Delta/\phi \in \mathfrak{D}$ then $\Delta \models \phi$.

PROOF. Since $\mathfrak{D}$ is recursively specified, this can be proved by induction. □

In working out the proof of this proposition, you will notice that the only rule whose validity depends on the fact that our information models are closed is M. However, even here 'depends' is too much said: M remains valid if the requirement of closedness is replaced by the much weaker requirement (iv') discussed in section III.1.1.

III.3.2. Completeness and decidability

This section contains a proof of the converse of proposition 14: the characterization of data logic given by $\mathfrak{D}$ is not only correct but also complete.

III.15. DEFINITION. Let Δ and Γ be sets of sentences.

(i) Δ is *consistent* iff there is some ϕ such that $\Delta/\phi \notin \mathfrak{D}$

(ii) Δ is *saturated* iff for any two sentences ϕ and ψ it holds that if $\phi \vee \psi \in \Delta$, then $\phi \in \Delta$ or $\psi \in \Delta$

(iii) Δ is a *theory within* Γ iff for every $\phi \in \Gamma$ it holds that if $\Delta/\phi \in \mathfrak{D}$ then $\phi \in \Delta$

(iv) Δ is *rich* iff Δ satisfies the following conditions.

(a) Δ is closed under subformulas: if $\phi \in \Delta$ and ψ is a subformula of ϕ then $\psi \in \Delta$

(b) Δ is closed under negations of proper subformulas: if $\phi \in \Delta$ and ψ is a subformula of ϕ such that $\psi \neq \phi$, then $\sim\psi \in \Delta$

(c) if $\phi \in \Delta$, and ϕ is atomic, then $(\textit{must}\ \phi \vee \sim \textit{must}\ \phi) \in \Delta$

(d) if $\phi \in \Delta$, and $\phi = \textit{must}\ \psi$ for some ψ, then $(\phi \vee \sim\phi) \in \Delta$. □

Readers acquainted with intuitionistic logic will have gathered from this definition that the completeness of $\mathfrak{D}$ is going to be proved along the lines of the completeness proofs for intuitionistic logic given by Aczel (1968) and Thomason (1968).

Note that the set of all sentences is rich. The only other interesting rich sets of sentences are the finite ones. And even these will not be interesting until we come to discuss the decidability of $\mathfrak{D}$.

In the following we will write 'Δ is a *CS-theory within* Γ' to abbreviate that Δ is a consistent and saturated theory within Γ. When Γ happens to be the set of all sentences we will suppress mention of Γ and just say that Δ is a CS-theory. In the completeness proof of $\mathfrak{D}$, CS-theories play very much the same role as maximal $\mathfrak{L}$-consistent sets of sentences did in the completeness proofs for the logics $\mathfrak{L}$ of part II.

III.16. LEMMA. Suppose $\Delta/\phi \notin \mathfrak{D}$. Let Γ be any set of sentences such that $\Delta \cup \{\phi\} \subset \Gamma$. Then there is an extension Δ' of Δ such that Δ' is a CS-theory within Γ and $\phi \notin \Delta$.

PROOF. Assuming that the language we are dealing with has countably many sentences[14)], we enumerate the sentences of Γ in such a way that each sentence in Γ occurs countably many times. Let $\sigma_1,\ldots,\sigma_n,\ldots$ be the the resulting enumeration. We define Δ' to be $\bigcup_{n\in\omega}\Delta_n$, where

(i) $\Delta_0 = \Delta$

(ii) each Δ_{n+1} is determined as follows

- if $\Delta_n/\sigma_n \notin \mathfrak{D}$ then $\Delta_{n+1} = \Delta_n$
- if $\Delta_n/\sigma_n \in \mathfrak{D}$, and σ_n is no disjunction, then $\Delta_{n+1} = \Delta_n \cup \{\sigma_n\}$
- if $\Delta_n/\sigma_n \in \mathfrak{D}$, and $\sigma_n = \psi \vee \chi$, then $\Delta_{n+1} = \Delta_n \cup \{\sigma_n,\psi\}$ if $\Delta_n,\psi/\phi \notin \mathfrak{D}$; otherwise $\Delta_{n+1} = \Delta_n \cup \{\sigma_n,\chi\}$

Notice

(a) If $\psi \in \Gamma$, and $\Delta'/\psi \in \mathfrak{D}$ then $\psi \in \Delta'$. Proof: Suppose $\Delta'/\psi \in \mathfrak{D}$. Given the compactness of $\mathfrak{D}$, we can find some n such that $\Delta_n/\psi \in \mathfrak{D}$ and $\sigma_n = \psi$. It follows that $\psi \in \Delta_{n+1}$. Hence $\psi \in \Delta'$.

(b) $\phi \notin \Delta'$. Proof: It is sufficient to show that their is no n such that $\Delta_n/\phi \in \mathfrak{D}$. That $\Delta_0/\phi \notin \mathfrak{D}$ is our starting point. Now suppose $\Delta_k/\phi \notin \mathfrak{D}$. We will show that $\Delta_{k+1}/\phi \notin \mathfrak{D}$. There are four cases (see above).

- $\Delta_{k+1} = \Delta_k$. This case is trivial.
- $\Delta_{k+1} = \Delta_k \cup \{\psi\}$ for some ψ such that $\Delta_k/\psi \in \mathfrak{D}$. Using the cut-rule it is easily seen that if Δ_{k+1}/ϕ were to belong to $\mathfrak{D}$, also Δ_k/ϕ would belong to $\mathfrak{D}$. Which, by the induction hypothesis, it does not.
- $\Delta_{k+1} = \Delta_k \cup \{\psi \vee \chi, \psi\}$, where $\Delta_k, \psi/\phi \notin \mathfrak{D}$. This case is obvious.
- $\Delta_{k+1} = \Delta_k \cup \{\psi \vee \chi, \chi\}$, where $\Delta_k/\psi \vee \chi \in \mathfrak{D}$, and $\Delta_k,\psi/\phi \in \mathfrak{D}$. Suppose $\Delta_k,\psi \vee \chi,\chi/\phi \in \mathfrak{D}$. Then $\Delta_k, \chi/\phi \in \mathfrak{D}$ (apply cut). Applying ETv, we see that $\Delta_k,\psi \vee \chi/\phi \in \mathfrak{D}$. Using the cut-rule, we get $\Delta_k/\phi \in \mathfrak{D}$, which contradicts the induction hypothesis.

(c) Δ' is saturated. Proof: Suppose $\psi \vee \chi \in \Delta'$. Then $\psi \vee \chi \in \Delta_k$ for some k. Consider any $n \geqq k$ such that $\sigma_n = \psi \vee \chi$. clearly $\psi \in \Delta_{n+1}$ or $\chi \in \Delta_{n+1}$. □

III.17. DEFINITION. Let Δ and Γ be sets of sentences.

$$\Delta \preccurlyeq \Gamma \text{ iff } \Delta^t \subset \Gamma^t$$ □

III.18. LEMMA. Let Γ be a rich set of sentences. Suppose Δ is a CS-theory within Γ. Then the following holds for any $\chi \in \Gamma$.

(i) if $\chi = \sim\sim\phi$ then $\chi \in \Delta$ iff $\phi \in \Delta$

(ii) if $\chi = (\phi \vee \psi)$ then $\chi \in \Delta$ iff $\phi \in \Delta$ or $\psi \in \Delta$

(iii) if $\chi = \sim(\phi \vee \psi)$ then $\chi \in \Delta$ iff $\sim\phi \in \Delta$ and $\sim\psi \in \Delta$

(iv) if $\chi = \mathit{must}\ \phi$ then $\chi \in \Delta$ iff there is no CS-theory Δ' within Γ such that $\Delta' \succcurlyeq \Delta$ and $\sim\phi \in \Delta'$

(v) if $\chi = \sim\mathit{must}\ \phi$ then $\chi \in \Delta$ iff there is some CS-theory Δ' within Γ such that $\Delta' \succcurlyeq \Delta$ and $\sim\phi \in \Delta'$

PROOF.

(i) holds in virtue of EF$\sim$ and IF$\sim$.

(ii) The one direction holds because Δ is saturated; the other in virtue of IT$\vee$.

(iii) holds in virtue of EF$\vee$ and IF$\vee$.

(iv) Suppose $\mathit{must}\ \phi \in \Delta$. Let Δ' be any CS-theory within Γ such that $\Delta' \succcurlyeq \Delta$. By the definition of $\preccurlyeq$, $\mathit{must}\ \phi \in \Delta'$. Hence $\sim\phi \notin \Delta'$, otherwise ET *must* would yield that Δ' is inconsistent.
For the converse, suppose $\mathit{must}\ \phi \notin \Delta$. Then $\Delta / \mathit{must}\ \phi \notin \mathfrak{D}$. Given IT *must*, we find $\Delta^t, \phi/\phi \notin \mathfrak{D}$. This means $\Delta^t \cup \{\sim\phi\}$ is consistent. By lemma 15 there is a CS-theory Δ' within Γ such that $\Delta^t \cup \{\sim\phi\} \subset \Delta'$. Clearly, $\Delta' \succcurlyeq \Delta$.

(v) Suppose $\sim\mathit{must}\ \phi \in \Delta$. Then $\mathit{must}\ \phi \notin \Delta$; otherwise EF *must* would give that Δ is inconsistent. Given (iv) it follows that there is a CS-theory Δ' within Γ such that $\Delta' \succcurlyeq \Delta$ and $\sim\phi \in \Delta'$.
For the converse, suppose $\sim\mathit{must}\ \phi \notin \Delta$. Since Γ is rich, $(\mathit{must}\ \phi \vee \sim\mathit{must}\ \phi) \in \Gamma$. In virtue of IF *must*, $(\mathit{must}\ \phi \vee \sim\mathit{must}\ \phi) \in \Delta$. Since Δ is saturated, and $\sim\mathit{must}\ \phi \notin \Delta$, it follows that $\mathit{must}\ \phi \in \Delta$. Given (iv) it is clear, then, that there is no CS-theory Δ' within Γ such that $\Delta' \succcurlyeq \Delta$ and $\sim\phi \in \Delta'$. □

III.19. DEFINITION. Let Γ be rich. Consider the triple
$M_\Gamma = \langle S, \leqq, V\rangle$ where
$S = \{\Delta \mid \Delta$ is a CS-theory within $\Gamma\}$;
for any two $\Delta, \Delta' \in S$, $\Delta \leqq \Delta'$ iff $\Delta \preccurlyeq \Delta'$;
V is the (partial) function which assigns to each $\Delta \in S$ and each atomic sentence ϕ the value $V_\Delta(\phi) = 1$ iff $\phi \in \Delta$, and the value $V_\Delta(\phi) = 0$ iff $\sim\phi \in \Delta$.

M_Γ is called the *canonical model* for Γ □

In view of proposition 13, we may rest assured that the function V is well-defined.

III.20. THEOREM. Let Γ be rich, and consider the canonical model $M_\Gamma = \langle S, \leqq, V\rangle$. M_Γ is a closed information model with the property that for every $\phi \in \Gamma$ and $\Delta \in S$,

$\Delta \models_M \phi$ iff $\phi \in \Delta$

$\Delta \;{}_M{=}\!| \; \phi$ iff $\sim\phi \in \Delta$

PROOF. First, we show that $\leqq$ is a partial order. From definition 17 it is clear that $\leqq$ is reflexive and transitive. That $\leqq$ is antisymmetric follows from the following claim:

if $\Delta, \Delta' \in S$, and $\Delta^t = \Delta'^t$ then $\Delta = \Delta'$

The proof of the claim is straightforward: show with induction to the complexity of ϕ that if $\phi \in \Gamma$ then (i) $\phi \in \Delta$ iff $\phi \in \Delta'$ and (ii) $\sim\phi \in \Delta$ iff $\sim\phi \in \Delta'$.

Next we show that each maximal chain in $\langle S, \leqq\rangle$ has a maximal element. Let S' be a maximal chain in $\langle S, \leqq\rangle$. Consider $\Delta' = \cup\{\Delta^t \mid \Delta \in S'\}$. Obviously, Δ' is consistent. By lemma 16 there is a CS-theory Δ'' within Γ such that $\Delta' \subset \Delta''$. Clearly, $\Delta'' \geqq \Delta$ for any $\Delta \in S'$. Clearly, $\Delta'' \in S'$, otherwise S' would not be a maximal chain. And it is also clear that there are no $\Delta''' \in S$ such that $\Delta'' \leqq \Delta'''$. In other words, Δ'' is a maximal element of S.

We must also show that if Δ is a maximal element of S V_Δ is total. Let $\phi \in \Gamma$ be atomic. Since Γ is rich *must* $\phi \vee \sim$*must* $\phi \in \Gamma$. Since $\Delta/($*must* $\phi \vee \sim$*must* $\phi) \in \mathfrak{D}$, and Δ is a CS-theory within Γ, *must* $\phi \vee \sim$*must* $\phi \in \Delta$. So, since Δ is saturated we have that either (i) *must* $\phi \in \Delta$; or

(ii) $\sim$*must* $\phi \in \Delta$. Let us first consider case (i). Using rule M, we see that if *must* $\phi \in \Delta$, then $\sim$*must* $\sim\phi \in \Delta$. By lemma 18(v), it follows that there must be a CS-theory $\Delta' \succcurlyeq \Delta$ within Γ such that $\phi \in \Delta'$. Since Δ is maximal, Δ' must coincide with Δ. Hence $\phi \in \Delta$. The proof of case (ii) is similar.

It remains to be proved by induction that for any $\phi \in \Gamma$ and any $\Delta \in S$, (i) $\Delta \models_M \phi$ iff $\phi \in \Delta$; and (ii) $\Delta \ {}_M\!\!=\!\!| \ \phi$ iff $\sim\phi \in \Delta$. Given lemma 18, the proof is straightforward. □

III.20. COROLLARY. If $\Delta \models \phi$ then $\Delta/\phi \in \mathfrak{D}$.

PROOF. Suppose $\Delta/\phi \notin \mathfrak{D}$. Let Γ be the set of all sentences. Γ is rich. Given lemma 16, there is a CS-theory Δ' within Γ such that $\Delta \subset \Delta'$ and $\phi \notin \Delta'$. Δ' is an information state in the canonical model $M_\Gamma = \langle S, \leq, V\rangle$ for Γ. Applying theorem 20 we see that $\Delta' \models_M \psi$ for all $\psi \in \Delta$, while $\Delta' \not\models_M \phi$. Hence $\Delta \not\models \phi$. □

III.21. COROLLARY. $\{\Delta/\phi \in \mathfrak{D} \mid \Delta \text{ is finite}\}$ is decidable.

PROOF. It suffices to show that if Δ is finite and $\Delta/\phi \notin \mathfrak{D}$, there is a *finite* model $M = \langle S, \leq, V\rangle$ such that for some $s \in S$, $s \models_M \psi$ for all $\psi \in \Delta$, while $s \not\models_M \phi$. From definition 15 it is clear that since $\Delta \cup \{\phi\}$ is finite, there is a *finite* rich Γ such that $\Delta \cup \{\phi\} \subset \Gamma$. It follows that the canonical model for Γ is finite. From the above it is clear that this model has all the properties desired. □

Johan van Benthem sketched an altogether different route to the main results reported here. In his (1984a) he shows that data logic is reducible to the modal logic $\mathbb{S}4.1$. Here $\mathbb{S}4.1$ is the weakest logic in the sense of definition II.6 extending $\mathbb{K}$ (see theorem II.44 and the subsequent remarks) and containing all arguments of the form

$\Box\phi/\phi$, $\Box\phi/\Box\Box\phi$, and $\Box\Diamond\phi/\Diamond\Box\phi$.

The proof is based on the insight that the data semantic truth- and falsity definitions can be encoded in the language of modal logic, provided that the underlying modal logic has enough 'locomotive power', as Van Benthem puts it, 'to drive the reduction.' For all data logical sentences ϕ, two modal translations, ϕ^+ and ϕ^-, are defined by simultaneous recursion:

- if ϕ is atomic then $\phi^+ = \Box\phi$ and $\phi^- = \Box{\sim}\phi$;
- $({\sim}\phi)^+ = \phi^-$; $({\sim}\phi)^- = \phi^+$;
- $(\phi \vee \psi)^+ = \phi^+ \vee \psi^+$; $(\phi \vee \psi)^- = {\sim}({\sim}\phi^- \vee {\sim}\psi^-)$
- $(\mathit{must}\ \phi)^+ = \Box{\sim}\phi^-$; $(\mathit{must}\ \phi)^- = \Diamond\phi^-$.

Now the result can be stated thus: let Δ/ϕ be an argument in the language of data logic. Then

Δ/ϕ is logically valid iff $\Delta^+/\phi^+ \in S4.1$.

Since S4.1. is known to be decidable, the decidability of data logic follows immediately. For the details of the proof the reader is referred to Van Benthem (1984a).

III.4. PROBLEMS AND PROSPECTS

III.4.1. Data lattices

As I have already mentioned, the definition of an information model $<S,\leqq,V>$ leaves many questions unanswered. To repeat, whenever $V_s(\phi) = 1$ for a given atomic sentence ϕ and information state s, we say 'ϕ is true on the basis of the evidence available at s', but the model $<S,\leqq,V>$ does not give us any clue as to what this evidence is. Another point is that hardly any constraints are placed on the number of information states that an information model should contain. Suppose, for example, that the model consists of two information states, s and s'; and that the atomic sentence ϕ is true on the basis of the evidence available at s, and false on the basis of the evidence at s'. Shouldn't there then be an information state s" such that s" < s and s" < s' while $V_{s''}(\phi)$ is undefined?

The data models defined below are more satisfactory in this respect. They are built not on a set of information states, but on a set of possible facts. These being the primitive entities of our models, I do not intend to say a great deal about their nature (they just are what possible worlds are made of). The only thing I will assume is that if two possible facts f and g can obtain together, this simultaneous occurrence of f and g qualifies as another possible fact. This fact is called *the combination of* f *and* g. Since we want to talk of the combination of f and g even if f and g cannot possibly hold together, we introduce as a technical convenience the so-called improper fact, and we stipulate that if f and g cannot obtain simultaneously, the combination of f and g amounts to this improper fact.

These considerations taken jointly give the set of possible facts the structure of a semi-lattice:

III.22. DEFINITION. A *data lattice* is a triple $\langle \mathcal{F}, \circ, 0 \rangle$ with the following properties:

(i) $0 \in \mathcal{F}$, $\mathcal{F} \sim \{0\} \neq \emptyset$;

(ii) $\circ$ is a binary operation on $\mathcal{F}$ such that

(a) $f \circ f = f$

(b) $f \circ g = g \circ f$

(c) $(f \circ g) \circ h = f \circ (g \circ h)$

(d) $0 \circ f = 0$ □

Explanation: The members of $\mathcal{F} \sim \{0\}$ are to be conceived of as the possible facts, '$f \circ g$' is to be read as 'the combination of f and g'. 0 is to be thought of as the improper fact. Given our informal remarks, it will be clear that the $\circ$-operation should have the properties laid down in (a)-(d). When $f \circ g = 0$, we shall often say that *f and g are incompatible* [15], and when $f \circ g = f$, we shall say that f *incorporates* g.

III.23. DEFINITION. Let $\langle \mathcal{F}, \circ, 0 \rangle$ be a data lattice. A possible world in $\langle \mathcal{F}, \circ, 0 \rangle$ is a subset $\mathcal{W}$ of $\mathcal{F}$ with the following properties:

(i) for every $f \in \mathcal{F}$, either $f \in \mathcal{W}$ or $g \in \mathcal{W}$ for some g such that $g \circ f = 0$;

(ii) for no $f_1, \ldots, f_n \in \mathcal{F}$, $f_1 \circ \ldots \circ f_n = 0$. □

A possible world is a rather peculiar set of possible facts: it is complete in the sense that if a given fact f does not obtain in it, some fact g incompatible with f obtains in it; and it is consistent in the sense that no incompatible facts obtain in it. Actually, possible worlds are so peculiar that one might wonder whether they exist at all. In other words, given any data lattice $\langle \mathcal{F}, \circ, 0 \rangle$, are there subsets of $\mathcal{F}$ meeting both the requirements (i) and (ii)? A well known theorem in lattice theory tells us that we may rest assured that this is the case. Before we can state the result, we need one more definition.

III.24. DEFINITION. Let $\langle \mathcal{F},\circ,0\rangle$ be a data lattice.
A *filter* in $\langle \mathcal{F},\circ,0\rangle$ is a subset of $\mathcal{F}$ meeting the following condition: for any $f,g \in \mathcal{F}$, $f,g \in \mathcal{G}$ iff $f \circ g \in \mathcal{G}$.
A filter $\mathcal{G}$ is *proper* iff $0 \notin \mathcal{G}$.
A proper filter $\mathcal{G}$ is *maximal* iff there is no proper filter $\mathcal{G}'$ such that $\mathcal{G} \subset \mathcal{G}'$ and $\mathcal{G} \neq \mathcal{G}'$. □

III.25. THEOREM. Let $\langle \mathcal{F},\circ,0\rangle$ be a data lattice.

(i) A subset $\mathcal{E}$ of $\mathcal{F}$ can be extended to a proper filter iff for every $f_1,\ldots,f_n \in \mathcal{D}$, $f_1 \circ \ldots \circ f_n \neq 0$;

(ii) Every proper filter can be extended to a maximal proper filter;

(iii) Every maximal proper filter is a possible world and *vice versa*.

PROOF. Omitted.[16] □

We are ready now to say what will be the analogue of an information state in this framework.

III.26. DEFINITION. Let $\langle \mathcal{F},\circ,0\rangle$ be a data lattice.
A (*possible*) *data set* in $\langle \mathcal{F},\circ,0\rangle$ is any subset $\mathcal{D}$ of $\mathcal{F}$ with the property that for any $f_1,\ldots,f_n \in \mathcal{D}$, $f_1 \circ \ldots \circ f_n \neq 0$. □

Given theorem 25, we see that any set of facts that might be obtained by investigating some possible world is a possible set of data. Thus, it would seem that a data set is just the right candidate for providing *the available evidence* at a possible information state. Think of a data set as the set of facts you are *acquainted* with in a given information state.

A few more observations before we pass on from ontology to semantics: notice that the theory of facts put forward here does not carry the metaphysical burden of many other theories. It is not assumed, for example, that there are facts of minimal complexity: any fact may incorporate other facts. Neither is it assumed that there are facts of maximal complexity: any fact may be incorporated by other facts. Furthermore, it is not assumed that there are *negative* facts: if a certain possible fact does not obtain in a certain possible world, then some possible fact g incompatible with f obtains in it, but there does not have to be some particular fact g incompatible with f which obtains in *every* possible world in which f does not obtain. And finally, it is not assumed that there are any *disjunctive* facts: suppose there are worlds in which f obtains and g does not obtain, as well as worlds in which g obtains and f does not. What we do *not* assume, then, is that there will be some fact h obtaining in exactly those worlds in which either f or g obtains.

In short, Occam's Razor has been of great help in developing definition 22.[18] Still, this does not mean that it might not be sensible to impose some additional constraints on the semilattice of possible facts. For example it would seem that much conceptual clarity is gained if we add the following condition to (a)-(d):

(e) if $f \circ g \neq f$ then there is some $h \in F$ such that $f \circ h \neq 0$ and $g \circ h = 0$

That is, if f does not incorporate g, then there is some fact h which is compatible with f but incompatible with g. In this way one ensures that if f does not incorporate g, there will be some possible world in which f obtains *without* g obtaining there as well. Or, to put it differently, in this way one ensures that any two *indistinguishable* facts, i.e. any two facts that obtain in exactly the same possible worlds, will be *equal*.[19]

III.26. DEFINITION. Let $\langle F, \circ, 0\rangle$ be a data lattice. An (atomic) *interpretation* I over $\langle F, \circ, 0\rangle$ is a function assigning some element $I(\phi)$ of F to each atomic sentence ϕ. $M = \langle F, \circ, 0, I\rangle$ is called a *data model*.

Informally, this definition can be put as follows: each atomic sentence describes a possible fact. Hence, in a way the definition offers a final clue to the question of what possible facts are: apparently, possible facts are things that can be described by the most elementary kind of sentences.[20]

Let $M = \langle F, \circ, 0, I \rangle$ be a model, D a data set in $\langle F, \circ, 0 \rangle$ and ϕ a sentence. In the following, '$D \models_M \phi$' will be short for 'ϕ is true (in M) on the basis of the evidence provided by D', and '$D\ {}_M\!\!=\!\mid \phi$' will be short for 'ϕ is false (in M) on the basis of the evidence provided by D'. And we stipulate that for atomic sentences the following is to hold:

- $D \models_M \phi$ iff $I(\phi) \in D$
- $D\ {}_M\!\!=\!\mid \phi$ iff for some $f \in D$, $f \circ I(\phi) = 0$

Given definition 2 it will be clear how to extend this to complex sentences. One example:

$D \models_M \phi \rightarrow \psi$ iff there is no data set $D' \supset D$ such that $D' \models_M \phi$ and $D'\ {}_M\!\!=\!\mid \psi$

$D\ {}_M\!\!=\!\mid \phi \rightarrow \psi$ iff there is some data set $D' \supset D$ such that $D' \models_M \phi$ and $D'\ {}_M\!\!=\!\mid \psi$

There are two reasons why I prefer data models to information models. The first is that data models show much more clearly than information models that data semantics can be understood as a correspondence theory of some sort. 'Truth' can still be equated with 'correspondence to the facts', albeit with the qualification that only the facts one is acquainted with matter. The same idea plays a key role in situation semantics, recently developed by Jon Barwise and John Perry (see Barwise and Perry (1984)). The main difference with situation semantics lies in the treatment of modality. Where here the notions of compatibility and incompatibility are introduced to make things work out properly, Barwise and Perry introduce so-called 'constraints' - a kind of higher order 'facts'.[21]

The second reason why I prefer data models to information models is that they allow for a smooth generalization of the present theory to other phenomena. In Landman (1984) a start has been made with a data semantic treatment of attitude reports, and this will be further developed in his forthcoming dissertation. In Groenendijk et.al. (forthcoming) it is shown that the ideas presented here can be fruitfully generalized so as to deal with problems of quantification and equality.

One final remark: It is obvious that the logic generated by data models is at least as strong as the logic generated by information models. It would be interesting to know whether they are the same. Though I conjecture that this is indeed the case, I will not attempt to prove it.

III.4.2. Complex conditionals

We must now face the fact that the theory of conditionals developed in the preceding pages does not in all respects work as well as may have been suggested. This becomes apparent if we take a closer look at arguments containing sentences of depth greater than one.

Let ϕ and ψ be atomic sentences. The argument $\phi \rightarrow$ *may* $\psi / \phi \rightarrow \psi$ is valid in data logic. In other words, the sentence

if the red marble is in box 1, the yellow marble may be in box 2

datalogically implies the sentence

if the red marble is in box 1, the yellow marble is in box 2

But clearly someone in information state 1 (see section 2.1) will be prepared to accept the first of these sentences, but not the second.

It is tempting to try to explain this counterexample away by saying that the first sentence is not really an implication, but has the form of a conjunction *may*(*red in 1* $\wedge$ *yellow in 2*). This is what Peter Geach (1976) advises us to do. Or we could follow Stalnaker (1976) and stipulate that the logical form of the first sentence is *may*(*red in 1* $\rightarrow$ *yellow in 2*).

We have already decided, however, to resist such temptations. (See section I.2.3.) Consequently, we must look for another way out.

Here are alternative clauses for *may*, *must* and $\rightarrow$.

$\mathcal{D} \models$ *must* ϕ iff for every $\mathcal{D}' \supset \mathcal{D}$ such that $\mathcal{D}' \mathrel{=\!|} \phi$, there is some $\mathcal{D}''$ with $\mathcal{D} \subset \mathcal{D}'' \subset \mathcal{D}'$ such that $\mathcal{D}'' \models \phi$

$\mathcal{D} \mathrel{=\!|}$ *may* ϕ iff $\mathcal{D} \not\models$ *must* ϕ

$\mathcal{D} \models$ *may* ϕ iff there is some $\mathcal{D}' \supset \mathcal{D}$ such that $\mathcal{D}' \models \phi$, while for no $\mathcal{D}''$ such that $\mathcal{D} \subset \mathcal{D}'' \subset \mathcal{D}'$, $\mathcal{D}'' \mathrel{=\!|} \phi$

$\mathcal{D} \mathrel{=\!|}$ *may* ϕ iff $\mathcal{D} \not\models$ *may* ϕ

$\mathcal{D} \models \phi \rightarrow \psi$ iff for every $\mathcal{D}' \supset \mathcal{D}$ such that $\mathcal{D}' \models \phi$ and $\mathcal{D}' \mathrel{=\!|} \psi$ there is some $\mathcal{D}''$ with $\mathcal{D} \subset \mathcal{D}'' \subset \mathcal{D}'$ such that $\mathcal{D}'' \models \phi$ and $\mathcal{D}'' \models \psi$

$\mathcal{D} \mathrel{=\!|} \phi \rightarrow \psi$ iff $\mathcal{D} \not\models \phi \rightarrow \psi$

It is not that definition 2 is wrong, it is just that its range of application is restricted. The clauses for $\phi \rightarrow \psi$, *may* ϕ, and *may* ψ given in definition 2 work well, but only for *stable* ϕ and ψ. Definition 2 could have been invented by someone who did not yet realize that there are unstable sentences, and that sentences can contain these unstable sentences as subsentences.

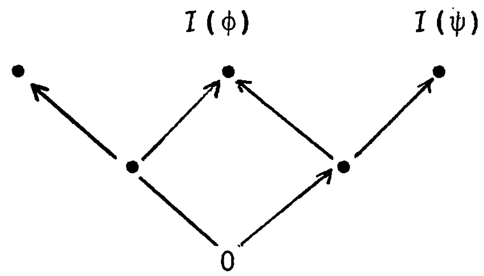
All sentences with depth less than 1 are stable. So we get

III.27. PROPOSITION. Let ϕ be a sentence with $d(\phi) \leqq 1$.
$\mathcal{D} \models \phi$ in the new sense of '$\models$' iff $\mathcal{D} \models \phi$ in the old sense of '$\models$';
$\mathcal{D} \mathrel{=\!|} \phi$ in the new sense of '$\mathrel{=\!|}$' iff $\mathcal{D} \mathrel{=\!|} \phi$ in the old sense of '$\mathrel{=\!|}$'.
PROOF. Left to the reader. □

In working out the proof of this proposition it will have become clear how the new clauses work. Definition 2 says that a sentence of the form $\phi \rightarrow \psi$ is true on the basis of the available data iff the possibility is excluded that upon further investigation ϕ turns out true and ψ false. This condition is somewhat weakened here: there might be extensions of the data on the basis of which ϕ is true and ψ is false, but $\phi \rightarrow \psi$ still counts as true if these data sets are as it were shielded by data sets on the basis of which both ϕ and ψ are true: you will never arrive at a situation where ϕ is true and ψ is false on the basis of the available data without arriving first at a situation where both ϕ and ψ are true.

Similar remarks apply to the other clauses. One more example: the new truth condition for *may* is somewhat stronger than the old one. For a sentence of the form *may* ϕ to be true it is not only necessary that there be extensions of the data on the basis of which ϕ is true; at least one of these extensions must be unshielded in the sense that the additional information can be acquired without ϕ's becoming false at any stage on the way.

Below a data model is pictured which shows that on this new account $\phi \rightarrow$ *may* ψ / $\phi \rightarrow \psi$ is no longer valid.



The dots represent the possible facts. Whenever f incorporates g an arrow is drawn from f to g. It is easy to see that in this model $\phi \rightarrow$ *may* ψ is true on the basis of the empty data set while $\phi \rightarrow \psi$ is false.

The above amendments solve some of the problems but unfortunately do not solve them all. For one thing it turns out that *may* $(\phi \rightarrow \psi)$ / $\phi \rightarrow \psi$ is valid on the new truth definitions and this may seem even worse than the validity of $\phi \rightarrow$ *may* / $\phi \rightarrow \psi$. Be this as it may, it is no reason for prefering the earlier truth definitions: if you think about it enough then it appears that these share the problem inherent in the new truth definitions which results in the validity of *may* $(\phi \rightarrow \psi)$ / $\phi \rightarrow \psi$, though without the same obviously counter-intuitive ramifications. I think that this problem is quite fundamental, and that it ultimately amounts to this: both truth definitions allow for the background knowledge which one employs to be partial; this is built into the whole idea of successive information states and the like. But the extent to which they do so is limited. Both truth definitions assume in a sense that ones knowledge of the *changes* which ones partial knowledge *could yet undergo* is complete. By freely quantifying over *all* possible extensions of partial information they assume that one, in evaluating conditionals, is in a position to take all of these possibilities into account, that one has a complete knowledge of the structure in question. This is of course not very true to real life.

Another result of this is the validity in both systems of the excluded middle for conditionals: they both give

$(\phi \rightarrow \psi) \vee \sim(\phi \rightarrow \psi)$.

Either $\phi \rightarrow \psi$ is true on the basis of the data or it is false. There is no room for question here. The problem is thus quite deep, and the modifications which would solve it go much further than just fiddling around with the truth definitions some more. The idea that a speaker may be only partially acquainted with the way the world is structured will have to be built into new truth conditions, and into new and presumably very complicated sorts of structures. Ample matter for further study.

III.4.3. Counterfactuals

Up till now, I have only dealt with indicative conditionals, so by way of conclusion I would like to comment on how counterfactuals can be treated within data semantics. The obvious first step is to try the same approach as earlier on in II.3.3. and give a formal version of Ramsey's suggestion. In the following truth definition ϕ and ψ are restricted to sentences of depth 0, to descriptive sentences, since as we saw above complex conditionals complicate matters enormously. '$\Rightarrow$' stands for counterfactual implication

III.28. DEFINITION. Let $\langle F, \circ, 0, I \rangle$ be a data model.

(i) D *admits* ϕ iff for some $D' \supset D$, $D' \models \phi$

(ii) $D \models \phi \Rightarrow \psi$ iff every ϕ-admitting $D' \subset D$ can be extended to a ϕ-admitting $D'' \subset D$ such that for no $D''' \supset D''$ it holds that $D''' \models \phi$ and $D''' \mathrel{=\!|} \psi$

This way, all of the schemes we had in $\mathfrak{P}$, the logic of premise functions, turn out valid. To sum up:

CI	$\emptyset \,/\, \phi \Rightarrow \phi$
CW	$\phi \Rightarrow \psi \,/\, \phi \Rightarrow (\psi \vee \chi)$
CC	$\phi \Rightarrow \psi,\ \phi \Rightarrow \chi \,/\, \phi \Rightarrow (\psi \wedge \chi)$
AD	$\phi \Rightarrow \chi,\ \psi \Rightarrow \chi \,/\, (\phi \vee \psi) \Rightarrow \chi$
ASC	$\phi \Rightarrow \psi,\ \phi \Rightarrow \chi \,/\, (\phi \wedge \chi) \Rightarrow \psi$

But in addition to these we also get:

ASP	$\phi \Rightarrow \psi,\ \sim(\phi \Rightarrow \sim\chi) \,/\, (\phi \wedge \chi) \Rightarrow \psi$
MP	$\phi,\ \phi \Rightarrow \psi \,/\, \psi$
CS	$\phi \wedge \psi \,/\, \phi \Rightarrow \psi$
while AS	$\phi \Rightarrow \psi \,/\, (\phi \wedge \chi) \Rightarrow \psi$

turns out invalid.

So it appears that the logic which results when Ramsey's advice is fitted into data semantics has a first degree fragment very similar to the logic which Lewis favours, which is based on an almost connected comparative similarity relation. In particular the validity of ASP is pleasing.

A few quite considerable problems do however remain. This approach leaves the problem of disjunctive antecedents unsolved, and it does not deal adequately with the puzzle presented by Paul Tichy.

Data semantics does suggest another and quite different way of dealing with counterfactuals. I think that it can best be illustrated by finally giving away which boxes contain which marbles. We already know (information state 2) the blue marble to be in box 1. What's new is that the yellow one is too, and that the red one is in box 2. And now a question: if the red marble had been in box 1, would the yellow one then have been in box 2? If you think that it would, then you are not evaluating counterfactuals as the previous definition would have you do so. What you are doing is this: you are going back to a time, or to be more precise to an information state, at which you did not yet know the red one to be in box 2. You go back to information state 2 and you

ask yourself 'Suppose I learned that the red marble is in box 1?'
In other words, what you are doing is going back to the *last* information state which you actually reached and at which you did not know the antecedent of the counterfactual you are being asked to evaluate to be true.

Turning counterfactuals into something indexical in this way will uncover more problems than it will solve. Still, I think it's worthwhile developing further.

NOTES TO PART III

1. The information models defined here closely resemble the Kripke models for intuitionistic logic. See Kripke (1965). Formally, the main difference with intuitionistic semantics lies in the treatment of negation. See Thomason (1969) for still another treatment of negation within this framework. It was Prof. N. Löb who suggested using these Kripke models instead of the data models of section 4.1.

2. See Kripke 1965: 98). I am ready to admit that the word 'information' as it occurs in the phrase 'information state' is not used in its ordinary sense. We will always be thinking of information as being correct information.

3. That one cannot impose condition (*) - negation by failure - when the information is incomplete is also relevant in computer science. This may become clear from following quotation which is taken from Barbara Partee's contribution to the *Report of Workshop on Information and Representation*, Washington D.C., 1985:

'Question-answering systems that are connected to data bases are typically designed so that when a yes-no question is asked, the machine tries to verify the corresponding assertion; if it succeeds, it answers 'Yes,' otherwise it answers 'No.' But sometimes it should in principle answer 'I don't know.' It turns out to be quite a difficult problem to design a system that knows when to say 'I don't know.' The 'closed world' assumption amounts to the dogmatic position 'If I don't know it, is isn't true.' Related problems arise for the semantics of negation in partial models, for the semantics of the programming language PROLOG which equates negation with failure, and in the analysis of propositional attitudes, where the distinctions among disbelief, agnosticism, and total absence of attitude can only be made in the context of powerful systems for reasoning about one's own beliefs. Awareness of one's own possible ignorance makes life infinitely more troublesome; but I believe it is one of the most significant among distinctively human attributes - it is, after all, what makes science possible, among other things. Computers that had some 'awareness' of their limitations could be not only more trustworthy, but certainly more user-friendly. This is an area with rich promise for both theoretical and practical payoff.'

4. Cases of irony and metaphor will perhaps be considered as counterexamples to this claim. But I think these phenomena are best explained as involving an *apparent* infringement of the maxim of quality. In short: since a literal interpretation of an ironical or metaphorical statement is out of the question, as it would immediately lead to the conclusion that the speaker is breaching the maxim of quality, the hearer tries to reinterpret the words of the speaker in such a way that they can as yet be reconciled with this maxim - the maxim of quality itself. Cases like these must be clearly distinguished from cases where the hearer ultimately concludes that a maxim - any maxim other than the maxim of quality - has *really* been overruled, albeit in a manner that can be reconciled with the supposition that at least the overall Cooperative Principle - but not the maxim in question - has been observed.

5. These critical remarks were made by Stanley Peters in his discussion of a talk I held at the Stanford Symposium on Conditionals.

6. See Hughes and Cresswell (1972: 47).

7. The statement 'He is a fine friend, if he is really telling all these lies' belongs - after reinterpretation - to

category 5. Compare footnote 4. (This example was brought to my notice by Richmond Thomason.)

8. Admittedly, this remark leaves a lot of questions concerning the example (k) open. For one thing, it is unclear why English speakers prefer (k) to the sentence *I paid back that fiver, if you don't remember*. Given our explanation for (i), one would expect things to be the other way around - as they are when one uses *in case* instead of *if*. (*In case you don't remember, I paid back that fiver* sounds better than *In case you remember, ...*) Only if the antecedent contains a negation one can safely say that it provides a condition under which the consequent would be informative.

9. Many of the examples discussed in this section have been taken from Lauerbach (1979). For a further discussion of, in particular, examples involving a maxim of politeness the reader is referred to pp. 240-250 of this book.

10. English allows both clause orders antecedent-consequent and consequent-antecedent. From the examples given so far, it appears that this is so even for conditionals that implicate the truth of their consequent. Notice, however, that one cannot overtly mark the consequent with *then* in some of these conditionals without affecting their original impact. This is particularly so for conditionals where 'if' is used as an opting out device, witness *If I may interrupt you, then you are wanted on the telephone*. In Dutch and German changing the word order in the consequent has the same effect: it seems obligatory to give the consequent the word order of a single main clause (finite verb second) when 'if' is used as an opting out device; while in all other cases with the antecedent preceding the consequent the verb of the consequent gets second position with respect to the antecedent clause and thus precedes the subject of the consequent. This means that the whole conditional construction is treated as a single main clause with the antecedent taking the front adverbial position.

11. In section 4.1 it will appear that this is to be understood literally.

12. It is instructive to compare this subdivision of true conditionals into normal laws, normal contingencies, degenerate laws and degenerate contingencies, which is given in the light of a uniform semantical treatment of conditionals, with the tripartition offered by Johnson-Laird (forthcoming), which is made as a (syntactic?) preamble to a pluriform treatment of conditionals. Those conditionals that would be classified by Johnson-Laird as true members of his first category would be classified as degenerate contingencies by me. The true members of Johnson-Laird's second category are all normal laws. And the true members of his third category are all what I would call normal contingencies.

13. See chapter 8 of Cooper (1978) for a load of other examples.

14. It is left to the reader to generalize the proof.

15. I want to stress that the improper fact is introduced merely as a technical convenience. In principle, one can dispense with it by taking a *partial* combination operation and calling two facts f and g incompatible iff the combination of f and g is not defined.

16. The proof is identical to the proof of the analogous theorem for Boolean Algebras. For details, see for example Bell and Slomson (1969: 13-15). It is worth noticing that (ii), though somewhat weaker than the Axiom of Choice, is independent of the axioms of Zermelo Fraenkel Set Theory. Therefore, its status as a mathematical truth is not as solidly based as these axioms - and so the existence of possible worlds becomes questionable is just as questionable as (ii).

17. The position on 'negative facts' taken here is not so different from Mr. Demos' position, which is discussed by

Bertrand Russell in *The Philosophy of Logical Atomism.* See the relevant chapter in Russell (1956).

18. Some may wonder whether it is possible to do without 'conjunctive' facts as well. It *is* possible but much of the simplicity and elegance of the data lattices gets lost.

19. See Landman (forthcoming) for arguments against this.

20. Admittedly, in the absense of a clear cut grammatical criterion to determine which English sentences count as most elementary, this remark is not very illuminating.

21. See Landman (1985) for a detailed comparison.

REFERENCES

ACZEL, P.H.G.: 1968, Saturated intuitionistic theories', in H.A. Schmidt, K. Schütte and H.J. Thiele (eds.), *Contributions to Mathematical Logic*, North Holland Publishing Company, Amsterdam.

ADAMS, E.: 1966, 'Probability and the logic of conditionals', in J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*, North Holland Publishing Company, Amsterdam.

ADAMS, E.: 1975, *The Logic of Conditionals. An Application of Probability to Deductive Logic*, Reidel, Dordrecht.

ANDERSON, A.R. and N.D. BELNAP: 1975, *Entailment, The Logic of Relevance and Necessity*, vol. I, Princeton University Press.

BACON, J.: 1971, 'The subjunctive conditional as relevant implication', *Philosophia 1*.

BARKER, J.A.: 1969, *A Formal Analysis of Conditionals*, Southern Illinois University Monographs, Carbondale.

BARTH, E.M.: 1974, 'Un timely remarks on the logic of "the modalities" in natural language', in C.H. Heidrich (ed.), *Semantics and Communication*, North Holland Publishing Company, Amsterdam.

BARWISE, J. and J. PERRY: 1983, *Situations and Attitudes*, MIT Press, Cambridge Mass.

BELL, J.L. and A.B. SLOMSON: 1969, *Models and Ultraproducts: An Introduction*, North Holland Publishing Company, Amsterdam.

van BENTHEM, J.F.A.K.: 1980, 'Some kinds of modal completeness', *Studia Logica 49*.

van BENTHEM, J.F.A.K.: 1984, ' Foundations of conditional logic', *Journal of Philosophical Logic 13*.

van BENTHEM, J.F.A.K.: 1984 a, *Partiality and Nonmonotonicity in Classical Logic*, Report No. CSLI - 84-12, Center for the Study of Language and Information, Stanford University.

van BENTHEM, J.F.A.K.: 1984 b, 'Correspondence Theory', in D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic, vol. II*, Reidel, Dordrecht.

BURGESS, J.P.: 'Quick completeness proofs for some logics of conditionals', *Notre Dame Journal of Formal Logic 22*.

CARROLL, L.: 1894, 'A logical paradox', *Mind 3*.

CHELLAS, B.F.: 1975, 'Basic conditional logic', *Journal of Philosophical Logic 4*.

CHISHOLM, R.M.: 1966, *Theory of Knowledge*, Prentice-Hall Englewood Cliffs N.J.

COOPER, W.S.: 1978, *Foundations of Logico-Linguistics; A Unified Theory of Information, Language and Logic*, Reidel, Dordrecht.

COPELAND, B.J.: 1978, *Entailment ;The Formalisation of Inference*, Thesis, submitted for the degree of Doctor of Philosophy in the University of Oxford.

ELLIS, B.: 1979, *Rational Belief Systems*, Basil Blackwell, Oxford.

FEYERABEND, P.K.`1970, 'Consolations for the specialist', in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*, Cambridge University Press.

FINE, K.: 1974, 'Logics containing K 4, part I', *Journal of Symbolic Logic 39*.

GAZDAR, G.: 1979, *Pragmatics;Implicature, Presupposition and Logical Form*, Academic Press, New York.

GEACH, P.I.: 1976, *Reason and Argument*, Basil Blackwell, Oxford.

GIBBARD, A.: 1981, 'Two Recent Theories of Conditionals', in W.L. Harper, R. Stalnaker and G. Pearce (eds.), *Ifs*, Reidel, Dordrecht.

GRICE, H.P.: 1975, 'Logic and Conversation', in P. Cole and J. Morgan (eds.), *Syntax and Semantics 3:Speech Acts*, Academic Press, New York.

GROENENDIJK, J and M. STOKHOF: 1975, 'Modality and Conversational Information', *Theoretical Linguistics 2*.

GROENENDIJK, J. and M. STOKHOF: 1982, 'Over logische vorm', *Tijdschrift voor Tekst- en Taalwetenschap, 2*.

GROENENDIJK, J., F. LANDMAN, M. STOKHOF and F. VELTMAN: forthcoming, 'Identity and Acquaintance in Data Semantics'.

HERZBERGER, H.G.: 1979, 'Counterfactuals and consistency', *Journal of Philosophy 76*.

HUGHES, G.E. and M.J. CRESSWELL: 1972, *An Introduction to Modal Logic*, Second edition, Methen and Co. Ltd., London.

JOHNSON-LAIRD, P.N.: 1985, 'Models of conditionals', in E. Traugott et al. (eds.), *On Conditionals*, Cambridge University Press.

KAMP, H.: 1984, 'A theory of truth and semantic representation', in J. Groenendijk, T.M.V. Janssen and M. Stokhof (eds.), *Truth, Interpretation and Information; selected Papers from the Third Amsterdam Colloquium*, Foris, Dordrecht.

KARTTUNEN, L.: 1972, 'Possible and must', in J. Kimball (ed.), *Syntax and Semantics 1*, Seminar Press, New York.

KRABBE, E.C.W.: 1978, 'Note on a completeness theorem in the theory of counterfactuals', *Journal of Philosophical Logic 7.*

KRATZER, A., 1977, 'What 'must' and 'can' must and can mean', *Linguistics and Philosophy 1.*

KRATZER, A.: 1979, 'Conditional necessity and possibility', in R. Bauerle, U. Egli and A. von Stechow (eds.), *Semantics from Different Points of View*, Springer-Verlag, Berlin.

KRATZER, A.: (1981), 'Partition and revision: the semantics of counterfactuals', *Journal of Philosophical Logic 10.*

KRIPKE, S.A.: 1965, 'Semantical analysis of intuitionistic logic I', in J.N. Crossley and M.A.E. Dummett (eds.), *Formal Systems and Recursive Functions*, North Holland Publishing Company, Amsterdam.

LANDMAN, F.: 1984, 'Data Semantics for Attitude Reports', in F. Landman and F. Veltman (eds.), *Varieties of Formal Semantics*, Foris, Dordrecht.

LANDMAN, F.: 1985, 'The realist theory of meaning', *Linguistics and Philosophy 8.*

LANDMAN, F.: forthcoming, *Towards a Theory of Information*, Foris, Dordrecht.

LAUERBACH, G.: 1979, *Form und Funktion englischer Konditionalsätze mit >if<*, Max Niemayer Verlag, Tübingen.

LEWIS, D.: 1973, *Counterfactuals*, Basil Blackwell, Oxford.

LEWIS, D.: 1974, 'Intensional logics without iterative axioms', *Journal of Philosophical Logic 3.*

LEWIS, D.: 1979, 'Probabilities of conditionals and conditional probabilities', *Philosophical Review 85.*

LEWIS, D: 1979, 'Counterfactual dependence and time's arrow', *Noûs 13.*

LEWIS, D.: 1981, 'Ordering semantics and premise semantics for counterfactuals', *Journal of Philosophical Logic 10.*

LOEWER, B: 1976, 'Counterfactuals with disjunctive antecedents', *Journal of Philosophy 73.*

LYONS, J.: 1977, *Semantics, vol. II,* Cambridge University Press.

McCAWLEY, J.D.: 1981, *Everything that Linguists have Always Wanted to Know about Logic (but were ashamed to ask)*, Basil Blackwell, Oxford.

McKAY, T. and P. van INWAGEN: 1977, 'Counterfactuals with disjunctive antecedents', *Philosophical Studies 31.*

NUTE, D.: 1978, 'An incompleteness theorem for conditional logic', *Notre Dame Journal of Formal Logic 19.*

NUTE, D.: 1980, *Topics in Conditional Logic,* Reidel, Dordrecht.

NUTE, D.: 1980 a, 'Conversational scorekeeping and conditionals', *Journal of Philosophical Logic 9.*

NUTE, D.: 1984, 'Conditional Logic', in D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic, vol. II,* Reidel, Dordrecht.

POLLOCK, J.: 1976, *Subjunctive Reasoning*, Reidel, Dordrecht.

POLLOCK, J.: 1981, 'A refined theory of counterfactuals', *Journal of Philosophical Logic 10.*

RAMSEY, F.P.: 1929, 'General Propositions and Causality', reprinted in F.P. Ramsey, *Foundations;Essays in Philosophy, Logic, Mathematics and Economics,* Routledge and Kegan Paul, London.

RESCHER, N.: 1964, *Hypothetical Reasoning*, North Holland Publishing Company, Amsterdam.

ROUTLEY, R. and R.K. MEYER: 1973, 'The semantics of entailment (I)', in H. Leblanc (ed.), *Truth, Syntax and Modality,* North Holland Publishing Company, Amsterdam.

RUSSELL, B.: 1956, *Logic and Knowledge; Essays 1901 - 1950*, edited by R.C. Marsh, George Allen and Unwin Ltd., London.

STALNAKER, R.: 1968, 'A theory of conditionals', in N. Rescher (ed.), *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, No. 2, Basil Blackwell, Oxford.

STALNAKER, R.: 1976, 'Indicative conditionals', in A. Kasher (ed.), *Language in Focus*, Reidel, Dordrecht.

STALNAKER. R.: 1976, 'Possible worlds', *Noûs 10*.

STALNAKER, R.: 1984, *Belief and Belief Ascription*, manuscript read at the 5th Amsterdam Colloquium.

STALNAKER, R. and R.H. THOMASON: 1970, 'A semantical analysis of conditional logic', *Theoria 36*.

STRAWSON, P.F.: 1952, *Introduction to Logical Theory*, Methuen, London.

THOMASON, R.H.: 1968, 'On the strong semantical completeness of the intuitionistic predicate calculus', *Journal of Symbolic Logic 33*.

THOMASON, R.H.: 1969, 'A semantical study of constructible falsity, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik 15*.

TICHY, P.: 1976, 'A counterexample to the Stalnaker-Lewis analysis of counterfactuals', *Philosophical Studies 29*.

TURNER, R.: 1981, 'Counterfactuals without possible worlds', *Journal of Philosophical Logic 10*.

VELTMAN, F.: 1976, 'Prejudices, presuppositions and the theory of counterfactuals', in J. Groenendijk and M. Stokhof (eds.), *Amsterdam Papers in Formal Grammer, vol. I*, Centrale Interfaculteit, Universiteit van Amsterdam.

VELTMAN, F.: 1984, 'Data Semantics', in J. Groenendijk, T.M.V. Janssen and M. Stokhof (eds.), *Truth, Interpretation and Information; Selected Papers from the Third Amsterdam Colloquium*, Foris, Dordrecht.

VELTMAN, F.: 1985, 'Data semantics and the pragmatics of indicative conditionals', in E. Traugott et al., *On Conditionals*, Cambridge University Press.

WARMBRÖD, K.: 1981, 'Counterfactuals and substitution of equivalent antecedents', *Journal of Philosophical Logic 10*.

SAMENVATTING

Wat is de logica van conditionele zinnen? Er zijn in de laatste decennia zoveel verschillende antwoorden op deze vraag gegeven dat met het simpelweg toevoegen van weer een nieuw antwoord aan de lange lijst van reeds bestaande, de problemen alleen maar groter lijken te worden. Daarom begint dit proefschrift met een methodologische verhandeling: hoe te kiezen tussen alternatieve logische theorieën? In die verhandeling bestrijd ik de opvatting dat logica een descriptieve wetenschap kan zijn. De vraag waarmee deze samenvatting begint is misleidend en dient vervangen te worden door de meer pragmatische vraag welke logica voor conditionele zinnen de beste is. (Zie ook stelling I).

Het tweede hoofdstuk van deel I van dit proefschrift is eveneens inleidend van aard. Aan de orde komen de scholen die zich met conditionele zinnen en hun logica bezig houden. Ze worden onderscheiden aan de hand van de explicatie van logische geldigheid die ze als uitgangspunt voor hun studie nemen. Aan de orde komen ook de voornaamste soorten van argumentatie waarmee men andermans standpunt bestrijdt en het eigen standpunt verdedigt. Ruwweg kunnen hier drie soorten van argumentatie onderscheiden worden - semantische, pragmatische en syntactische argumentatie - maar als complicerende factor geldt dat niet iedereen hetzelfde onder deze begrippen verstaat. De verschillende opvattingen over logische geldigheid geven aanleiding tot evenzovele opvattingen over wat semantiek is - of moet zijn. En hoewel men het er in het algemeen over eens is dat de pragmatische component de vorm moet aannemen van een conversatietheorie à la Grice zijn er grote meningsverschillen over de inhoud ervan - dat kan ook niet anders aangezien zo'n conversatietheorie

een semantische theorie vooronderstelt. Wat de syntactische component betreft is er in de genoemde scholen nauwelijks sprake van gedegen theorievorming. Daarin brengt dit proefschrift trouwens geen verandering. Waarin het wel van de meeste andere studies op het gebied van de conditionele logica verschilt is dat er zo min mogelijk een toevlucht genomen wordt tot ad hoc argumentatie op syntactisch gebied.

In deel II worden de semantische theorieën die ontwikkeld zijn door de belangrijkste van bovengenoemde scholen nader bestudeerd. Het betreft hier de theorieën ontwikkeld binnen het raamwerk van de zogenaamde mogelijke werelden semantiek. Na een beschrijving van dit raamwerk volgt een vergelijking, zowel filosofisch als mathematisch, van de verschillende analyses van counterfactuals die men heeft voorgesteld.
Aan de orde komen theorieën gesteld in termen van selectiefuncties, (met name die van Stalnaker (1968)), theorieën die gebruik maken van een comparatieve similariteitsrelatie tussen mogelijke werelden (zoals in Lewis (1973)), en theorieën die gebruik maken van premissefuncties (Veltman (1976) en Kratzer (1979)). Middels een aantal representatie stellingen worden deze theorieën zoveel mogelijk onder één noemer gebracht. Deze stellingen bieden ook de sleutel tot de oplossing van een aantal open problemen van meer technische aard.

Eén van de bevindingen in deel II is dat bij het gebruik en de interpretatie van conditionele zinnen de achtergrondinformatie waarover sprekers en luisteraars beschikken een cruciale rol speelt. De theorie gesteld in termen van premisse functies kan daarbij gezien worden als een eerste poging om deze rol te expliciteren. Deze theorie blijkt echter niet in alle opzichten even bevredigend. Om wel tot zo'n theorie te komen worden in deel III enkele van de inzichten opgedaan in deel II geradicaliseerd. Waar in deel II de rol van achtergrondinformatie beperkt blijft tot die van een contextuele factor, die de waarheidswaarde van *sommige* soorten van zinnen mede bepaalt, krijgt in deel III achtergrondinformatie de rol van een determinerende factor bij de interpretatie van *alle* soorten van zinnen. Niet waarheid *simpliciter* maar waarheid *op grond van de beschikbare informatie* wordt daarbij tot fundamentele semantische notie gemaakt. Door deze perspectiefverschuiving komen een aantal tot nog toe onopgemerkt gebleven aspecten van conditionele zinnen aan het licht. Zo blijkt bijvoorbeeld dat in tegenstelling tot puur descriptieve zinnen conditionele zinnen niet stabiel zijn: ze kunnen bij uitbreiding van de beschikbare informatie van waarheidswaarde veranderen. Een aantal logische eigenaardigheden van conditionele zinnen blijken direct op deze instabiliteit teruggevoerd te kunnen worden (zie ook stelling III).

Het grootste gedeelte van deel III wordt besteed aan een vergelijking van deze nieuwe theorie met de bestaande. In dat kader wordt ook ruime aandacht besteed aan de repercussies van deze semantische aanpak voor de pragmatiek (zie ook stelling I). Op het wiskundige vlak beperkt het onderzoek zich tot het ontwikkelen van een deductiesysteem en het bewijs van de volledigheid daarvan.

Stellingen
behorende bij het proefschrift *Logics for Conditionals*
van F.J.M.M. Veltman.

I

Of de voorspellingen van een logische theorie goed, minder goed, of zelfs slecht aansluiten bij de intuïties van de taalgebruikers is een kwestie van ondergeschikt belang. Waar het om gaat is of die theorie zo goed gemotiveerd is dat de taalgebruikers zich erdoor laten leiden en corrigeren, juist bij gelegenheden dat de voorspellingen van de theorie niet - nog niet - aansluiten bij hun intuïtieve verwachtingen.

[Dit proefschrift, hoofdstuk I.1]

II

Men moet van een semantiek niet verlangen dat ze het predikaat 'waar' slechts aan die conditionele zinnen toekent waarbij sprake is van een 'causaal' verband tussen antecedent en consequent. Zo'n semantiek zou geen ruimte bieden om pragmatisch te verklaren hoe het komt dat de volgende zinnen waar kunnen zijn ondanks de afwezigheid van een dergelijk verband.

- Als het morgen niet gaat regenen dan gaat het gieten.
- Als ik ergens een hekel aan heb dan is het aan vroeg opstaan.
- Als u me nog nodig mocht hebben dan ben ik bereikbaar op toestel 4564.

[Dit proefschrift, hoofdstuk III.2]

III

Noem een zin ϕ *stabiel* als ϕ de volgende eigenschappen heeft: (i) als ϕ waar is op grond van een verzameling gegevens dan blijft ϕ waar bij uitbreiding van die gegevens; (ii) als ϕ onwaar is op grond van een verzameling gegevens dan blijft ϕ onwaar bij uitbreiding van die gegevens.

In tegenstelling tot zuiver descriptieve zinnen zijn conditionele zinnen, zinnen die beginnen met 'misschien' en zinnen die beginnen met 'vast en zeker' niet stabiel. Het grillige logische gedrag van deze zinnen is gegrond in deze instabiliteit.

[Dit proefschrift, hoofdstuk III.1]

IV

Noem een kardinaalgetal σ te verwaarlozen ten opzichte van kardinaalgetal τ als i) τ eindig is en σ=0, of ii) τ oneindig is en σ < τ. Noem een conditionele zin waar als het aantal gevallen waarin het antecedent waar is maar het consequent onwaar te verwaarlozen is ten opzichte van het aantal gevallen waarin antecedent en consequent beide waar zijn. De logica voor conditionele zinnen die door deze waarheidsdefinitie wordt gegenereerd, valt samen met David Lewis' systeem V.

[J. v. Benthem: 'Foundations of Conditional Logic', *Journal of Philosophical Logic*, 13, 1984]

V

Het belang van de begrippen 'exhaustiviteit', 'rigiditeit' en 'definietheid' voor een theorie over de vraag-antwoord relatie blijkt duidelijker wanneer de stellingen 12 en 35 uit hoofdstuk V.4. van het proefschrift van Jeroen Groenendijk en Martin Stokhof als volgt worden versterkt en samengevat: Een term α is in een bepaalde kontekst c exhaustief, rigide en definiet dan en slechts dan als α in c een compleet antwoord is op elke vraag van de vorm 'Wat zijn de β's?' waarop α geen contradictoir antwoord is.

[J. Groenendijk and M. Stokhof: *Studies in the Semantics of Questions and the Pragmatics of Answers*. Dissertatie, Universiteit van Amsterdam, 1984]

VI

Een bevredigende oplossing van de Sorites Paradox dient gebaseerd te zijn op een semantiek voor vage predikaten die het begrijpelijk maakt dat van een verzameling zinnen waarin vage predikaten optreden, elke zin,afzonderlijk beschouwd, waar kan zijn terwijl die verzameling als geheel toch incoherent is.
Hans Kamps oplossing is van deze aard, maar zijn semantiek voor vage predikaten kan nog aanzienlijk vereenvoudigd worden.

[H. Kamp: 'The Paradox of the Heap', in
U. Mönnich (ed.) *Aspects of Philosophical Logic*,
Reidel, Dordrecht, 1981]

VII

Stelling 6.124 uit Wittgensteins Tractatus kan gegeneraliseerd worden:
Jede Logik beschreibt das Gerüst einer Welt, oder vielmehr sie stellt es dar.

[L. Wittgenstein: *Tractatus logico-philosophicus*]

VIII

Het zal de kwaliteit van het universitair onderwijs ten goede komen als meer docenten er een gewoonte van maken studenten niet met hun onderzoeksresultaten lastig te vallen voordat ook docenten aan andere universiteiten deze daarvoor interessant genoeg vinden.

IX

Zelfs bij het knikkeren gaat het niet om het spel.