GRAHAM PRIEST

PARACONSISTENT LOGIC

Indeed, even at this stage, I predict a time when there will be mathematical investigations of calculi containing contradictions, and people will actually be proud of having emancipated themselves from 'consistency'. Ludwig Wittgenstein, 1930.[1]

1 INTRODUCTION

Paraconsistent logics are those which permit inference from inconsistent information in a non-trivial fashion. Their articulation and investigation is a relatively recent phenomenon, even by the standards of modern logic. (For example, there was no article on them in the first edition of the *Handbook*.) The area has grown so rapidly, though, that a comprehensive survey is already impossible. The aim of this article is to spell out the basic ideas and some applications. Paraconsist logic has interest for philosophers, mathematicians and computer scientists. As befits the *Handbook*, I will concentrate on those aspects of the subject that are likely to be of more interest to philosopher-logicians. The subject also raises many important philosophical issues. However, here I shall tread over these very lightly—except in the last section, where I shall tread over them lightly.

I will start in part 2 by explaining the nature of, and motivation for, the subject. Part 3 gives a brief history of it. The next three parts explain the standard systems of paraconsistent logic; part 4 explains the basic ideas, and how, in particular, negation is treated; parts 5 and 6 discuss how this basic apparatus is extended to handle conditionals and quantifiers, respectively. In part 7 we look at how a paraconsistent logic may handle various other sorts of machinery, including modal operators and probability. The next two parts discuss the applications of paraconsistent logic to some important theories; part 8 concerns set theory and semantics; part 9, arithmetic. The final part of the essay, 10, provides a brief discussion of some central philosophical aspects of paraconsistency.

In writing an essay of this nature, there is a decision to be made as to how much detail to include concerning proofs. It is certainly necessary to include many proofs, since an understanding of them is essential for anything other than a relatively modest grasp of the subject. On the other hand, to prove everything in full would not only make the essay extremely long, but distract from more important issues. I hope that I have struck a happy *via media.*

[1] Wittgenstein [1975], p. 332.

D. Gabbay and F. Guenthner (eds.),
Handbook of Philosophical Logic, Volume 6, 287–393.

Where proofs are given, the basic definitions and constructions are spelled out, and the harder parts of the proof worked. Routine details are usually left to the reader to check, even where this leaves a considerable amount of work to be done. In many places, particularly where the material is a dead end for the purposes of this essay, and is easily available elsewhere, I have not given proofs at all, but simply references. Those for whom a modest grasp of the subject is sufficient may, I think, skip all proofs entirely.

Paraconsistent logic is strongly connected with many other branches of logic. I have tried, in this essay, not to duplicate material to be found in other chapters of this *Handbook*, and especially, the chapter on Relevant Logic. At several points I therefore defer to these. There is no section of this essay entitled 'Further Reading'. I have preferred to indicate in the text where further reading appropriate to any particular topic may be found.[2]

2 DEFINITION AND MOTIVATION

2.1 Definition

The major motivation behind paraconsistent logic has always been the thought that in certain circumstances we may be in a situation where our information or theory is inconsistent, and yet where we are required to draw inferences in a sensible fashion. Let $\vdash$ be any relationship of logical consequence. Call it *explosive* if it satisfies the condition that for all α and β, $\{\alpha, \neg\alpha\} \vdash \beta$, *ex contradictione quodlibet* (ECQ). (In future I will omit set braces in this context.) Both classical and intuitionist logics are explosive. Clearly, if $\vdash$ is explosive it is not a sensible inference relation in an inconsistent context, for applying it gives rise to *triviality*: everything. Thus, a minimal condition for a suitable inference relation in this context is that it not be explosive. Such inference relationships (and the logics that have them) have come to be called *paraconsistent.*[3]

Paraconsistency, so defined, is something of a minimal condition for a logic to be used as envisaged; and there are logics that are paraconsistent but not really appropriate for the use. For example, Johansson's minimal logic is paraconsistent, but satisfies $\alpha, \neg\alpha \vdash \neg\beta$. One might therefore attempt a stronger constraint on the definition of 'paraconsistent', such as: for no syntactically definable class of sentences (e.g., negated sentences), Σ, do

[2]The most useful general reference is Priest *et al.* [1989] (though this is already a little dated). That book also contains a bibliography of paraconsistency up to about the mid-1980s.

[3]The word was coined by Miró Quesada at the Third Latin American Symposium on Mathematical Logic, in 1976. Note that a paraconsistent logic need not itself have an inconsistent set of logical truths: most do not. But there are some that do, e.g., any logic produced by adding the connexivist principle $\neg(\alpha \rightarrow \neg\alpha)$ to a relevant logic at least as strong as B. See Mortensen [1984].

we have $\alpha, \neg\alpha \vdash \sigma$, for all $\sigma \in \Sigma$. This seems too strong, however. In many logics, $\alpha, \neg\alpha \vdash \beta$, for every logical truth, β. If the logic is decidable, then there is a clear sense in which the set of logical truths is syntactically characterisable. Yet such logics would still be acceptable for many paraconsistent purposes. Hence, this definition would seem to be too strong.[4]

In his [1974], da Costa suggests another couple of natural constraints on a paraconsistent logic, of a rather different nature. One is to the effect that the logic should not contain $\neg(\alpha \wedge \neg\alpha)$ as a logical truth. The rationale for this is not spelled out. However, I take it that the idea is that if one has information that contains α and $\neg\alpha$ one does not want to have a logical truth that contradicts this. Why not though? Since one is not ruling out inconsistency *a priori*, there would seem to be nothing *a priori* against this (though maybe for particular applications one would not want the situation to arise). As a general condition, then, it seems too strong. And certainly a number of the logics that we will consider have $\neg(\alpha \wedge \neg\alpha)$ as a logical truth.

Another of the constraints that da Costa suggests is to the effect that the logic should contain as much of classical—or at least intuitionist—logic, as does not interfere with its paraconsistent nature. The condition is somewhat vague, though its intent is clear enough; and again, it is too strong. It assumes that a paraconsistent logician must have no objection to other aspects of classical or intuitionist logic, and this is clearly not true. For example, a relevant logician might well object to paradoxes of implication, such as $\alpha \to (\beta \to \alpha)$.[5]

As an aside, let me clarify the relationship between relevant logics and paraconsistent logics. The motivating concern of relevant logic is somewhat different from that of paraconsistency, namely to avoid paradoxes of the conditional. Thus, one may take a relevant (propositional) logic to be one such that if $\alpha \to \beta$ is a logical truth then α and β share a propositional parameter. The interests of relevant and paraconsistent logics clearly converge at many points. Relevant logics and paraconsistent logics are not coextensive, however. There are many paraconsistent logics that are not relevant, as we shall see. The relationship the other way is more complex, since there are different ways of using a relevant logic to define a consequence relation. A natural way is to say that $\alpha \vdash \beta$ iff $\alpha \to \beta$ is a logical truth. Such a consequence relation is clearly paraconsistent. Another is to define logical consequence as deducibility, defined in the standard way, using some set of axioms and rules for the relevant logic. Such a consequence relation may, but need not, be relevant. For example, Ackermann's original formulation of E contained the rule γ: if $\vdash \alpha$ and $\vdash \neg\alpha \vee \beta$ then $\vdash \beta$. This gives explo-

[4]Further attempts to tighten up the definition of paraconsistency along these lines can be found in Batens [1980] (in the definition of 'A-destructive', p. 201, clause (i) should read $\nvdash_L A$), and Urbas [1990].

[5]Indeed, it is just this principle that ruins minimal logic for serious paraconsistent purposes. For α and $\alpha \to \perp$ (i.e., $\neg\alpha$) give $\perp$, and the principle then gives $\beta \to \perp$.

sion by an argument often called the 'Lewis Independent Argument', that we will meet in a moment.

Anyway, and to return from the digression: the definition of paraconsistency given here is weaker than sufficient to *guarantee* sensible application in inconsistent contexts; but an elegant stronger definition is not at hand, and since the one in question has become standard, I will use it to define the contents of this essay.

2.2 Inconsistency and Dialetheism

Numerous examples of inconsistent information/theories from which one might want to draw inferences in a controlled way have been offered by paraconsistent logicians. For example:

1. information in a computer data base;
2. various scientific theories;
3. constitutions and other legal documents;
4. descriptions of fictional (and other non-existent) objects;
5. descriptions of counterfactual situations.

The first of these is fairly obvious. As an example of the second, consider, e.g., Bohr's theory of the atom, which required bound electrons both to radiate energy (by Maxwell's equations) and not to (since they do not spiral inwards towards the nucleus). As an example of the third, just consider a constitution that gives persons of kind A the right to do something, x, and forbids persons of kind B from doing x. Suppose, then, that a person in both categories turns up. (We may assume that it had never occurred to the legislators that there might be such a person.) In the fourth case, the information (in, say, a novel or a myth) characterises an object, and turns out—deliberately or otherwise—to be inconsistent. To illustrate the fifth, suppose, for example, that we need to compute the truth of the conditional: if you were to square the circle, I would give you all my money. Applying the Ramsey-test, we see what follows from the antecedent (which is logically impossible), together with appropriate background assumptions. (And I would *not* give you all my money!)[6]

There is no suggestion here that in every case one must remain content with the inconsistent information in question. One might well like to remove

[6]Many of these examples are discussed further in Priest *et al.* [1989], ch. 18. The Bohr case is discussed in Brown [1993]. Another kind of example that is sometimes cited is the information provided by witnesses at a trial. I find this less persuasive. It seems to me that the relevant information here is all of the form: witness x says so and so. (That a witness is lying, or making an honest mistake, is always a possibility to be taken into account.) And any collection of statements of this form is quite consistent.

some of the inconsistent information in the data base; reject or revise the scientific theory; change the law to eliminate the inconsistency. But this is not possible in all of the cases given, e.g., for counterfactual conditionals with impossible antecedents. And even where it is, this not only may take time; it is often not clear how to do so satisfactorily. (The matter is certainly not algorithmic.) While we figure out how to do it, we may still be in a situation where inference is necessary, perhaps for practical ends, e.g., so that we can act on the information in the data base; or manipulate some piece of scientific technology; or make decisions of law (on other than an obviously inconsistent case). Moreover, since there is no decision procedure for consistency, there is no guarantee that any revision will achieve consistency. We cannot, therefore, be sure that we have succeeded. (This is particularly important in the case of the data base, where the deductions go on "behind our back", and the need to revise may never become apparent.)

In cases of this kind, then, even though we may not, ideally, be satisfied with the inconsistent information, it may be desirable—indeed, practically necessary—to use a paraconsistent logic. Moreover, we know that many scientific theories are false; they may still be important because they make correct predications in most, or even all, cases; they may be good *approximations* to the truth, and so on. These points remain in force, even if the theories in question contain contradictions, and so are (thought to be) false for logical reasons. Of course, this is not so if the theories are trivial; but that's the whole point of using a paraconsistent logic.

One can thus subscribe to the use of paraconsistent logics for some purposes without believing that inconsistent information or theories may be *true*. The view that some *are* true has come to be called *dialetheism*, a dialetheia being a true contradiction.[7] If the truth about some subject is dialetheic then, clearly, a paraconsistent logic needs to be employed in reasoning about that subject. (I take it to be uncontentious that the set of truths is not trivial. Why this is so, especially once one has accepted dialetheism is, however, a substantial question.)

Examples of situations that may give rise to dialetheias, and that have been proposed, are of several kinds, including:

1. certain kinds of moral and legal dilemmas;
2. borderline cases of vague predicates;
3. states of change.

Thus, one may suppose, in the legal example mentioned before, that a person who is A and B both has and has not the right to do x; or that in

[7]The term was coined by Priest and Routley in 1981. See Priest *et al.* [1989], p. xx. Note that some writers prefer 'dialethism'.

a case of light drizzle it both is and is not raining; or that at the instant a moving object comes to rest, it both is and is not in motion.[8]

The most frequent and, arguably, most persuasive examples of dialetheias that have been given are the paradoxes of self-reference, such as the Liar Paradox and Russell's Paradox. What we have in such cases, are apparently sound arguments resulting in contradictions. There are many suggestions as to what is wrong with such arguments, but none of them is entirely happy. Indeed, in the case of the semantic paradoxes there is not (even after 2,000 years) any consensus concerning the most plausible way to go. This gives the thought that the arguments are, after all, sound, its appeal.[9]

Naturally, all the examples cited in this section are contestable. I will return to the issue of possible objections in the last part of this essay.

3 A BRIEF HISTORY OF PARACONSISTENT LOGIC

3.1 The Law of Non-contradiction and Paraconsistency

During the history of Western Philosophy, there have been a number of figures who deliberately endorsed inconsistent views.[10] The earliest were some Presocratics, including Heraklitus. In the middle ages, some Neo-Platonists, such as Nicholas of Cusa, endorsed contradictory views. In the modern period, the most notable advocate of inconsistent views was Hegel.[11] These figures are relatively isolated, however. It is something of an understatement to say that the dominant orthodoxy in Western Philosophy has been strongly hostile to inconsistency.[12] Consistency has been taken to be pivotal to a number of fundamental notions, such as truth and rational belief. This antipathy to contradiction is, historically, due in large part to Aristo-

[8]Many of these examples are discussed further in Priest *et al.* [1989], ch. 18, 2.2. A discussion of transition states and legal dialetheias can be found in chs. 11 and 13 of Priest [1987]. Moral dilemmas are also discussed in Routley and Routley [1989]. The dialetheic nature of vagueness is advocated in Peña [1989]. It has also been suggested that some contradictions in the Hegel/Marx tradition are dialetheic. For a discussion of this, see Priest [1989a].

[9]For further discussion, see Routley [1979] and Priest [1987], chs. 1-3.

[10]And nearly every great philosopher has unwittingly endorsed inconsistent views.

[11]In each case, there is, of course, some—though, I would argue, misguided—possibility for exegetical attempts to render the views consistent. Other modern philosophers whose thought also appears to endorse inconsistency are Meinong and the later Wittgenstein. In their cases there is more scope for exegetical evasion. For further discussion on all these matters, see Priest *et al* [1989], chs. 1, 2.

[12]Eastern philosophy has been notably less so—though there is, again, room for exegetical debate. The most natural interpretation of Jaina philosophy has them endorsing inconsistent positions. And major Buddhist logicians of the stature of Nārgārjuna held that it was quite possible for statements to be both true and false. Significant elemements of inconsistency can also be found in Chinese philosophy. For further discussion of all this, see Priest *et al* [1989], ch. 1, sect. 2.

tle's defense of the Law of Non-contradiction in the *Metaphysics*.[13] Given this situation, it may therefore be surprising that the orthodoxy against paraconsistency is a relatively recent phenomenon.

3.2 Paraconsistency Before the Twentieth Century

The major account of validity until this century was, of course, Aristotelian Syllogistic. Now, consider any sentences of the Syllogistic E and I forms; for example, 'No women are white' and 'Some women are white'. These are contradictories. But the inference from them to, e.g., 'All cows are black', is not a valid syllogism. Syllogistic is not, therefore, explosive: it is paraconsistent.

It might be suggested that it is more appropriate to look for explosion in accounts of propositional inference. Here the story is more complex, but the conclusion is similar. Aristotle had no elaborated account of propositional inference. However, there are comments that bear on the matter scattered through the *Organon*, and they have a distinctly paraconsistent flavour. For example, in the *Prior Analytics* (57^b3), Aristotle states that contradictories cannot both entail the same thing. It would seem to follow that Aristotle did not endorse at least one of (in modern notation) $\alpha \wedge \neg\alpha \vdash \alpha$ and $\alpha \wedge \neg\alpha \vdash \neg\alpha$. For contraposing (a move that Aristotle endorses immediately before), we obtain $\alpha \vdash \neg(\alpha \wedge \neg\alpha)$ and $\neg\alpha \vdash \neg(\alpha \wedge \neg\alpha)$. Hence, not everything can follow from a contradiction. In fact, there are reasons to suppose that Aristotle held a view of negation according to which the negation of any claim cancels that claim out. A contradiction has, therefore, no content, and entails nothing. This view of negation (which would now be called 'connexivist') was endorsed by a number of subsequent logicians (notably Abelard) well into the late middle ages.[14]

A theory of propositional inference was worked out much more thoroughly by Stoic logicians, and the explosive nature of their theories is more plausible for the following reason. There is a famous argument for ECQ, often called the Lewis (independent) argument, after C. I. Lewis. This goes (in natural deduction form) as follows:

$$\dfrac{\alpha \qquad \dfrac{\neg\alpha}{\neg\alpha \vee \beta}}{\beta}$$

[13]Book Γ, ch 4. The historical success of this defence is, however, out of all proportion to its intellectual weight. See Priest [1998e].

[14]Much of this and the rest of the material in this subsection is documented in Sylvan [2000], ch. 4. The discussion there is carried out in terms of the conditional, though it is equally applicable to the consequence relation.

The argument uses just two principles (three if you include the transitivity of deducibility): Addition ($\alpha \vdash \alpha \vee \beta$) and the Disjunctive Syllogism ($\alpha, \neg\alpha \vee \beta \vdash \beta$). As we shall see in due course, the Disjunctive Syllogism (DS) has, unsurprisingly, been rejected by most paraconsistent logicians. Now Stoic logicians endorsed just this principle. The explosive nature of their logic would therefore seem a good bet. Despite this, it probably was not: there is reason to suppose that their disjunction was an intensional one that required some kind of connection between α and β for the truth of $\alpha \vee \beta$. If this is the case, Addition fails in general, as does the Lewis argument.

It is not known who discovered the Lewis argument. Martin [1985] conjectures that it was William of Soissons in the 12th Century. (It was certainly known to, and endorsed by, some later logicians, such as Scotus and Buridan.) At any rate, William was a member of a group of logicians called the Parvipontanians, who were known not only for living by a small bridge, but for defending ECQ. This group may therefore herald the arrival of explosion on the philosophical stage. Whether or not this is so, after this time, some logicians endorsed explosion, some rejected it, different orthodoxies ruling at different times and different places (though, possibly, the explosive view was more common). One group of logicians who rejected it is notable, since they very much prefigure modern paraconsistent logicians. This is the Cologne School of the late 15th Century, who argued against the DS on the ground that if you start by *assuming* that α and $\neg\alpha$, then you cannot appeal to α to rule out $\neg\alpha$ as the DS manifestly does.

Notoriously, logic made little progress between the end of the Middle Ages and the start of the third great period in logic, towards the end of the 19th Century. With the work of logicians such as Boole and Frege, we see the mathematical articulation of an explosive logical theory that has come to be know, entirely inappropriately, as 'classical logic'. Though, in its early years, many objected to its explosive features, it has achieved a hegemony (though never a universality) in the logical community, in a (historically) very brief space of time. Whether this is because the truth was definitively and transparently revealed, or because at the time it was the only game in town, history will tell.

3.3 The Twentieth Century

A feature of paraconsistent logic this century is that the idea appears to have occurred independently to many different people, at different times and places, working in ignorance of each other, and often motivated by somewhat different considerations. Some, notably, for example, da Costa, have been motivated by the idea that inconsistent theories might be of intrinsic importance. Others, notably the early relevant logicians, were motivated simply by the idea that explosion, as a property of entailment, is just too counter-intuitive.[15]

The earliest paraconsistent logics (that I am aware of) were given by two Russians. The first of these was Vasil'év. Starting about 1910, Vasil'év proposed a modified Aristotelian syllogistic, according to which there is a new form: S is both P and not P. How, exactly, this form was to be interpreted is contentious, though, a problem exacerbated by the fact that he was not in a position to employ the techniques of modern logic. This is not true of the second logician, Orlov, who, in 1929, gave the first axiomatisation of the relevant logic R. Sadly, the work of neither Vasil'év nor Orlov made any impact at the time.[16]

An important figure who did have a good deal of influence was the Polish logician and philosopher Łukasiewicz. Partly influenced by Meinong's account of impossible objects, Łukasiewicz clearly envisaged the construction of paraconsistent logics in his seminal 1910 critique of Aristotle on the Law of Non-contradiction.[17] And it was his erstwhile student, Jaśkowski, who, in 1948, produced the first non-adjunctive paraconsistent logic.[18]

Paraconsistent logics were again, independently, proposed in South America in doctoral dissertations by Asenjo (1954, Argentina) and da Costa (1963, Brazil). Asenjo proposed the first many-valued paraconsistent logic. Da Costa gave axiom systems for a certain family of paraconsistent logics (the C systems), and produced the first quantified paraconsistent logic. Many co-workers, such as Arruda and Loparić, joined da Costa in the next 20 years, to produce an active school of paraconsistent logicians at Campinas (and later São Paulo). They developed non-truth-functional semantics for the C systems, and articulated the subject in various other ways; this included "rediscovering" Vasil'év, taking up the work of Jaśkowski, and formulating various other paraconsistent systems.[19]

Guided by considerations of relevance, an entirely different approach to paraconsistency was proposed in England by Smiley in [1959], who articulated the first filter logic. Starting at about the same time, and drawing on the earlier work of Ackermann and Church, Anderson and Belnap in the USA proposed a number of relevant paraconsistent logics of a different kind. A research school quickly grew up around them in Pittsburgh, which included co-workers such as Meyer and Dunn.[20] The algebraic semantics

[15]The later Wittgenstein was also sympathetic to paraconsistency for various reasons, though he never articulated a paraconsistent logic. See, e.g., Marconi [1984].

[16]On Vasil'év see Priest *et al.* [1989], ch. 3, 2.2 and Arruda [1977]. On Orlov, see Anderson *et al.* [1992], p. xvii.

[17]A synopsis of this is published in English in Łukasiewicz [1971].

[18]For a discussion of Łukasiewicz and Jaśkowski, see Priest *et al.* [1989], ch. 3, 2.1, 2.3. Jaśkowski's work is translated into English in his [1969].

[19]Discussion and bibliography can be found in Priest *et al.* [1989], 5.6. The most accessible introduction to Asenjo's work is his [1966], and to da Costa's is his [1974]. Da Costa and Marconi [1989] reports much of the work of da Costa and his co-workers.

[20]The work of this school is recorded in Anderson and Belnap [1975], and Anderson *et*

for relevant logics, in particular, was inaugurated largely by Dunn's 1966 doctoral thesis.[21]

Investigation of things paraconsistent in Australia took off in the early 1970s with the discovery of world (intensional) semantics for negation by R. Routley (now Sylvan) and V. Routley (now Plumwood). This was developed into an intensional semantics for the Anderson/Belnap logics—and many others—by Routley,[22] Meyer (now in Australia), and a school that developed around them in Canberra, which included workers such as Brady and Mortensen. These semantics made the paraconsistent aspects of relevant logics plain.[23] Later in the 1970s the cudgel for dialetheism was taken up by Priest (now Priest) and Routley.[24]

By the mid-1970s the paraconsistent movement was a fully international one, with workers in all countries cooperating (though not necessarily agreeing!), and with logicians working in numerous countries other than the ones already mentioned, including Belgium, Bulgaria, Canada and Italy. Some feel for the state of the subject at the end of the 70s can be obtained from Priest *et al.* [1989].[25] The rest, as they say, is not history.

4 BASIC TECHNIQUES OF PARACONSISTENT LOGICS

An understanding of most paraconsistent logics can be obtained by looking at the strategies employed in virtue of which ECQ fails. There are many techniques for achieving this end. In this part, I will describe the most fundamental. In the process, we will meet dozens of different systems of paraconsistent logic, often constructed along very different lines. It is therefore necessary to have some common medium for comparison. I have chosen to make this semantics, and will specify systems in terms of these. (I would warn straight away though, that many of the systems we will meet appeared first in proof theoretic terms. Indeed, some of the authors of these systems—e.g. Tennant—would privilege proof theory over semantics.) When I give details of corresponding proof theories, I will use the sort of proof theory (natural deduction, sequent calculus, or axiomatic) that seems most natural for the logic.

Because paraconsistency concerns only negation essentially, we can see the essentials of paraconsistent logics in languages with very little logical

al. [1992].

[21]See Anderson and Belnap [1975], and also the article on Relevant Logic in this volume of the *Handbook.*

[22]Whenever the name 'Routley' is used without initial in this essay, it refers to Sylvan.

[23]The work of this group is most accessible in Routley *et al.* [1982].

[24]See, e.g., Routley [1979]. Priest's early work on the area is most accessible in Priest [1987].

[25]Despite the date, all the work in the collection was finished by 1980. A number of papers produced at the same time, that were not included in this, were published in a special issue of *Studia Logica* on paraconsistent logics (**43** (1984), nos. 1 & 2).

apparatus. In this part, we will be concerned with a propositional language whose only connectives are negation, $\neg$, conjunction, $\wedge$, and disjunction, $\vee$. I will use lower case Roman letters, starting with p, for propositional parameters, lower case Greeks, starting with α, for arbitrary formulas, and upper case Greeks for sets of formulas. I will use $\models_C$, $\models_I$, and $\models_{S5}$ for the consequence relations of classical logic, intuitionist logic and $S5$, respectively, and $\models$ for the semantic consequence relation of whichever paraconsistent logic happens to be the topic of discussion. If a proof theory is involved, I will use $\vdash$ for the corresponding notion of deducibility.

4.1 Filtration

One of the simplest ways to prevent explosion is to filter it, and any other undesirables, out. Consider, for simplicity, the one-premise case. (Finite sets of premises can always be reduced to this by conjoining.) Let $F(\alpha, \beta)$ be any relationship between formulas. Define an inference from α to β to be *prevalid* iff $\alpha \models_C \beta$ and $F(\alpha, \beta)$. The thought here is that for an inference to be correct, something more than classical truth-preservation is required, e.g., some *connection* between premise and conclusion. This is expressed by F. Usually, prevalidity is too weak as a notion of validity, since, in general, it is not closed under uniform substitution, and this is normally taken to be a desideratum for any notion of validity. However, closure can be ensured if we define an inference to be valid iff it is a uniform substitution instance of a prevalid inference.

What inferences are valid depends, of course, entirely on the filter, F. One that naturally and obviously gives rise to a paraconsistent logic is: $F(\alpha, \beta)$ iff α and β share a propositional parameter. (This collapses the notions of validity and prevalidity, since if α and β share a propositional parameter, so do uniform substitution instances thereof.) This logic is not a very interesting paraconsistent one, however, since, as is clear, $p \wedge \neg p \models \alpha$ where α is any formula containing the parameter p.[26]

A different filter, proposed by Smiley [1959] is: $F(\alpha, \beta)$ iff α is not a (classical) contradiction and β is not a (classical) tautology.[27] (Note that, according to this definition, $\alpha \wedge \neg\alpha \ / \ \alpha$ is not prevalid, but it is valid, since it is an instance of $p \wedge q \ / \ p$.) It is easy to see that on this account $p \wedge \neg p$ does not entail q. The major notable feature of filter logics is that, in general, transitivity of deducibility breaks down.[28] For example, using

[26]A stronger filter is one to the effect that *all* the variables of the premise occur in the conclusion. This gives rise to a logic in the family of analytic implications. On this family, see Anderson and Belnap [1975], sect. 29.6.

[27]Filters of a related kind were also suggested by Geach and von Wright. See Anderson and Belnap [1975], sect. 20.1.

[28]Though it need not. First Degree Entailment, where transitivity holds, can be seen as a filter logic. See Dunn [1980].

Smiley's filter, it is easy to see that $p \wedge \neg p \models p \wedge (\neg p \vee q)$, $p \wedge (\neg p \vee q) \models q$, but $p \wedge \neg p \not\models q$.

One of the most interesting filter logics, given by Tennant [1984], is obtained by generalising Smiley's approach. Let Π and Σ be sets of sentences, and let '$\Pi \models_C \Sigma$' be understood in the natural way (every classical evaluation that makes every member of Π true makes some member of Σ true). Define the inference from Π to Σ to be prevalid iff: $\Pi \models_C \Sigma$ and for no proper subsets of Π and Σ, Π' and Σ', respectively, do we have $\Pi' \models_C \Sigma'$. Validity is then defined by closing under substitution as before. In this account, a valid inference is one which is classically valid, and minimally so: there is no "noise" amongst premise and conclusion set.[29]

Tennant's $\models$ is obviously non-monotonic (that is, adding extra premises may invalidate an inference). It also has the following property: if $\Pi \models_C \Sigma$, then there are subsets of Π and Σ, Π' and Σ', respectively, such that $\Pi' \models \Sigma'$. For if $\Pi \models_C \Sigma$, we can simply throw out premises and/or conclusions until this is no longer true; the result is a prevalid, and so valid, inference. In particular, if $\Pi \models_C \alpha$ then for some $\Pi' \subseteq \Pi$, $\Pi' \models \alpha$ or $\Pi' \models \phi$. In the first case, α follows validly from *part* of Π; in the second, part of Π can be shown to be inconsistent by valid reasoning.

Filtration can also be applied proof theoretically: we start with classical proofs and throw out those that do not satisfy some specific criteria. Tennant's logic can be characterised proof-theoretically in just this way. For finite premises and conclusions, the valid inferences are exactly those that are provable in the Gentzen sequent calculus for classical logic, but which do not use the structural rules of dilution (thinning) and cut. Specifically, consider the sequent calculus whose basic sequents are of the form $\alpha : \alpha$, and whose rules are as follows. (Π_1, Π_2 means $\Pi_1 \cup \Pi_2$; similarly, Π, α means $\Pi \cup \{\alpha\}$, and if something of this form occurs as a premise of a rule, it is to be understood that $\alpha \notin \Pi$).

$$\frac{\Pi, \alpha : \Delta}{\Pi : \Delta, \neg\alpha} \qquad \frac{\Pi : \Delta, \alpha}{\Pi, \neg\alpha : \Delta}$$

$$\frac{\Pi, \alpha : \Delta}{\Pi, \alpha \wedge \beta : \Delta} \qquad \frac{\Pi, \beta : \Delta}{\Pi, \alpha \wedge \beta : \Delta} \qquad \frac{\Pi_1 : \Delta_1, \alpha \quad \Pi_2 : \Delta_2, \beta}{\Pi_1, \Pi_2 : \Delta_1, \Delta_2, \alpha \wedge \beta}$$

$$\frac{\Pi : \Delta, \alpha}{\Pi : \Delta, \alpha \vee \beta} \qquad \frac{\Pi : \Delta, \beta}{\Pi : \Delta, \alpha \vee \beta} \qquad \frac{\Pi_1, \alpha : \Delta_1 \quad \Pi_2, \beta : \Delta_2}{\Pi_1, \Pi_2, \alpha \vee \beta : \Delta_1, \Delta_2}$$

[29]The restriction of Tennant's approach to the one-premise, one-conclusion, case obviously gives Smiley's account. Smiley himself, handles the multiple-premise case, simply by conjoining. As Tennant points out ([1984], p. 199), this generates a different account from his. It is not difficult to check that $p \vee q, \neg(p \vee q) \not\models p \wedge q$ for Smiley. (The conjoined antecedent is a contradiction; and any inference of which the conjoined form is a substitution instance is not classically valid.) But it is valid for Tennant, since it is a substitution instance of $p \vee q, r \vee s, \neg(t \vee q), \neg(r \vee u) \models p \wedge s$.

Then we have $\Pi \models \Sigma$ iff the sequent $\Pi : \Sigma$ is provable. For the proof see Tennant [1984].[30]

Tennant's account of inference seems to capture very nicely what one might call the 'essential core' of classical inference. As an inference engine to be applied to inconsistent information/theories, it could not be applied in the obvious way, however. This is because information is often heavily redundant. For example, for Tennant's $\models$, we do not have $p, \neg p \vee q, q \models q$. Yet given the information in the premises, it would certainly seem that we are entitled to infer q. Presumably, then, we would take α to follow from Σ iff for some $\Sigma' \subseteq \Sigma$, $\Sigma' \models \alpha$.[31] If we do this then more than transitivity fails; so does Adjunction. For $\neg p, p \vee q \models q$ and $\neg q, p \vee q \models p$, hence both p and q follow from $\{\neg p, p \vee q, \neg q\} = \Sigma$. But for no subset of Σ, Σ', do we have $\Sigma' \models p \wedge q$. ($\Sigma \models \phi$, and if Σ' is a proper subset of Σ, $\Sigma' \not\models_C p \wedge q$.) In this respect, Tennant's approach is similar to the next one that we will look at.

4.2 *Non-adjunction*

All the other approaches that we will consider, except the last (algebraic logics) accept validity as defined simply in terms of model-preservation. Thus, given some notion of interpretation, call it a *model* of a sentence if the sentence holds in the interpretation; an interpretation is a model of a set of sentences if it is a model of every member of the set; and an inference is valid iff every model of the premises is a model of the conclusion. In particular, then, if explosion is to be avoided, it must be possible to have models for contradictions, which are not models of everything. Where the differences in the following approaches lie is in what counts as an interpretation, and what counts as holding in it.

For the next approach, an interpretation, I, is a Kripke interpretation of some modal logic, say $S5$, employing the usual truth conditions. Each world in an interpretation may be thought of as the world according to some party in a debate or discussion. This gives the approach its common name, ***discussive*** (or ***discursive***) logic. I is a (discursive) model of sentence α iff α holds at some world in I, i.e., $\Diamond\alpha$ holds in the model. Thus, $\Sigma \models \alpha$ iff α holds, discursively, in every discursive model of Σ, i.e., iff $\Diamond\Sigma \models_{S5} \Diamond\alpha$, where $\Diamond\Sigma$ is $\{\Diamond\alpha; \alpha \in \Sigma\}$. This approach is that of Jaśkowski [1969].[32] It is clear that discussive logic is paraconsistent, since we may have $\Diamond\alpha$ and $\Diamond\neg\alpha$

[30] The proof theory can be given a filtered natural deduction form too. Essentially, classical deductions that have a certain "normal form" pass through the filter. See Tennant [1980].

[31] Though if we do this, symmetry suggests that we should take Π to follow from Σ iff for some $\Sigma' \subseteq \Sigma$ and $\Pi' \subseteq \Pi$, $\Sigma' \vdash \Pi'$; in this case paraconsistency is lost since $\alpha, \neg\alpha \vdash \phi$.

[32] Popper also seems to have had a similar idea in 1948. See his [1963], p. 321.

in an $S5$ interpretation, without having $\Diamond\beta$. For similar reasons, Adjunction ($\alpha, \beta \models \alpha \wedge \beta$) fails. It should be noted, however, that $\alpha \wedge \neg\alpha \models \beta$, so the logic is not paraconsistent for conjoined contradictions.

A closely related approach can be found in Rescher and Brandom [1979]. They define validity as truth preservation in all worlds, but they augment the worlds of standard modal logic by inconsistent and complete worlds, constructed using operators $\dot{\cup}$ and $\dot{\cap}$. Specifically, worlds are constructed recursively from standard worlds as follows. If W is a set of worlds, $\dot{\cup}W$ is a world such that α is true in $\dot{\cup}W$ iff for some w in W, α is true in w; and $\dot{\cap}W$ is a world such that α is true in $\dot{\cap}W$ iff for all w in W, α is true in w. As is intuitively clear, inconsistent worlds just provide another way of expressing what holds in a Jaśkowski interpretation. Incomplete worlds appear more novel, but, in fact, add nothing. For if truth fails to be preserved in one of these, it fails to be preserved in one of the ordinary worlds which go into making it up. These ideas can be recast to show that the semantics of Rescher and Brandom, and of Jaśkowski are inter-translatable, and deliver the same notion of validity.[33]

A notable feature of discussive logic is that $\Sigma \models \alpha$ iff for some $\beta \in \Sigma$, $\beta \models_C \alpha$. (The proof from right to left is obvious. From left to right, suppose that for every $\beta \in \Sigma$, $\beta \not\models_C \alpha$. Let w_β be a classical world where β holds but α does not. If we take the interpretation whose worlds are $\{w_\beta; \beta \in \Sigma\}$ this is a counter-model for $\Sigma \models \alpha$.) Thus, single-premise discussive inference is classical, and there is no essentially multiple-premise inference. One way to avoid the second of these features is to add an appropriate conditional connective. We will look at this later. Another way is to allow a certain amount of conjoining of premises. The question is how to do this in a controlled way so that explosion does not arise.

One suggestion, made by Rescher and Manor, is, in effect, to allow conjoining up to maximal consistency.[34] Given a set of premises, Σ, a maximally consistent subset (mcs) is any consistent subset, Σ', such that if

[33]Proof: Suppose that, discursively, $\Sigma \not\models \alpha$. Then there is an interpretation such that for each $\sigma \in \Sigma$, there is some world, w_σ, such that σ is true in w_σ, but α is not true in w_σ. Let $w = \dot{\cup}\{w_\sigma; \sigma \in \Sigma\}$, then w is a Rescher/Brandom counter-model. Conversely, suppose that $\Sigma \not\models \alpha$ for Rescher and Brandom. Then there is some world such that for every $\sigma \in \Sigma$, σ is true at w, but α is not. We show that there is a Jaśkowski counter-model. The result is proved by recusion on the construction of Rescher/Brandom worlds. If w is a standard world, the result is clear. So suppose that $w = \dot{\cap}W$, where the result holds for all members of W. By definition, for every $z \in W$, and every $\sigma \in \Sigma$, σ is true in z, but for some z, α is not true in z. Consider that z. This is a Rescher/Brandom counter-model to the inference. Hence, the result holds by recursion hypothesis. Alternatively, suppose that $w = \dot{\cup}W$, where the result holds for all members of W. By definition, for every $\sigma \in \Sigma$, there is some $w_\sigma \in W$, such that σ is true in w_σ, but α is not. By recursion, there must be a Jaśkowski countermodel for the inference σ/α. σ is true at some world in this, but α is not. If we form the collection of worlds for all such σ, this then gives us a Jaśkowski counter-model to the original inference.

[34]Rescher and Manor [1970-1]. This takes off from the earlier work of Rescher [1964].

$\alpha \in \Sigma - \Sigma'$, $\Sigma' \cup \{\alpha\}$ is inconsistent. We can now say that α follows from Σ iff for some mcs Σ', $\Sigma' \models_C \alpha$. In possible world terms, we can rephrase this as follows. Let us say that an interpretation, I, *respects* Σ iff for every mcs Σ', there is a world, w, in I such that Σ' is true in w. Then it is not difficult to see that this policy is a variant of discussive logic: $\Sigma \models \alpha$ iff α holds discussively in every interpretation that respects Σ. (If α follows classically from some Σ', then it holds in every discussive interpretation that respects Σ. Conversely, suppose that it follows from no Σ'. Then for each Σ' choose a world $w_{\Sigma'}$ where Σ' is true, but α is not. The interpretation containing all such $w_{\Sigma'}$ is a countermodel.)

This policy is certainly stronger than simple discussive consequence. For example, it gives: $p, q \models p \wedge q$. In fact, if Σ is (classically) consistent then every classical consequence of α is a consequence. But it is still non-adjunctive: $p, \neg p \not\models p \wedge \neg p$.[35]

A slightly different way of proceeding is provided by Schotch and Jennings.[36] Given a finite set, Σ, a *partition* is any family of disjoint sets, each of which is classically consistent, and whose union is Σ. The *level* of Σ, $l(\Sigma)$, is the least n such that Σ can be partitioned into n sets (or, conventionally, ∞ if there is no such n). $\Sigma \models \alpha$ iff $l(\Sigma) = \infty$ or, $l(\Sigma) = n$ and for any partition of Σ of size n, $\{\Sigma_i; 1 \leq i \leq n\}$, there is an i such that $\Sigma_i \models_C \alpha$. As with the previous approach, this definition can be converted into discussive terms, by taking our models to be those that respect the premise set. But this time, an interpretation respects Σ iff for some partition of the level of Σ, $\{\Sigma_i; 1 \leq i \leq n\}$, and every i, there is a world in the interpretation where Σ_i is true.

Leaving aside the fact that Schotch and Jennings consider only finite premise sets, one difference between their approach and the previous one concerns the consequences of sets, Σ, with single inconsistent members. Such sets have no partitions, and so explode for Schotch and Jennings. They still have mcss though (e.g., ϕ), and so do not explode for Rescher and Manor. If Σ has no single inconsistent member then Schotch and Jennings' consequence relation is included in that of Rescher and Manor. For if $\{\Sigma_i; i \in n\}$ is a partition of the premises, Σ, and for some i, $\Sigma_i \models_C \alpha$, then Σ_i can be extended to an mcs of Σ, and this classically entails α. The converse is not true, however. Let $\Sigma = \{p, \neg p, q, r\}$. This has two mcss, $\{p, q, r\}$ and $\{\neg p, q, r\}$. Hence, for Rescher and Manor, $q \wedge r$ follows. But Σ has level 2, and one partition is $\{\{p, q\}, \{\neg p, r\}\}$. Neither of these classically entails $q \wedge r$, so this does not follow for Schotch and Jennings (which seems wrong, intuitively).[37]

[35] Rescher and Manor also formulate a weaker policy of inference. α follows from Σ iff for *all* mcs Σ', $\Sigma' \models_C \alpha$. This logic is clearly adjunctive.

[36] See their [1980], where they also discuss appropriate proof theories and modal connections.

[37] The same example shows that Schotch and Jennings' $\models$, unlike Rescher and Manor's,

Despite the differences, Schotch and Jennings' approach shares with that of Rescher and Manor the following features: for consistent sets, consequence coincides with classical consequence; Adjunction fails. For Schotch and Jennings, like Jaśkowski, $\alpha \wedge \neg\alpha$ explodes. For Rescher and Manor, it has no consequences (other than tautologies). The logics that we will look at in subsequent sections are more discriminating concerning conjoined contradictions.

4.3 Interlude: Henkin Constructions

Before we move on to look at the other basic approaches to paraconsistent logic, I want to isolate a construction that we will have many occasions to use. In a standard Henkin proof for the completeness of an explosive logic, we construct a maximally consistent set of sentences, and use this to define an evaluation. In the construction of the set, we keep something out of it by putting its negation in. As might be expected, these techniques do not work in paraconsistent logic; but they can be generalised to do so. What plays the role of a maximally consistent set in a paraconsistent logic is a *prime theory*, where a set of sentences, Σ, is a theory iff it is closed under deducibility; and it is prime iff $\alpha \vee \beta \in \Sigma \Rightarrow (\alpha \in \Sigma \text{ or } \beta \in \Sigma)$. To keep something out in the construction of a prime theory, we have to exclude it explicitly. I now show how.

Assume that the proof theory is to be given in natural deduction terms. For definiteness I adopt the notational conventions of Prawitz [1965].[38] Consider the following rules for conjunction and disjunction:

$$\begin{array}{lll}
\vee I & \dfrac{\alpha \quad \beta}{\alpha \wedge \beta} & \\[2ex]
\wedge E & \dfrac{\alpha \wedge \beta}{\alpha} & \dfrac{\alpha \wedge \beta}{\beta} \\[2ex]
\vee I & \dfrac{\alpha}{\alpha \vee \beta} & \dfrac{\beta}{\alpha \vee \beta}
\end{array}$$

is non-monotonic, since we have $q, r \models q \wedge r$.

[38]In particular, something of the form:

$$\begin{array}{c} \alpha \\ \vdots \\ \beta \end{array}$$

in a rule indicates a subproof with α as one assumption—though there may be others—and conclusion β. If α is overlined, this means that the application of the rule discharges it.

$$\vee E \quad \dfrac{\alpha \vee \beta \qquad \begin{matrix} \overline{\alpha} \\ \vdots \\ \gamma \end{matrix} \qquad \begin{matrix} \overline{\beta} \\ \vdots \\ \gamma \end{matrix}}{\gamma}$$

Let $\vdash$ be any proof theory that includes these rules. Write $\Sigma \vdash \Pi$ to mean that there are members of Π, $\pi_1, ..., \pi_n$, such that $\Sigma \vdash \pi_1 \vee ... \vee \pi_n$. Then if $\Sigma \nvdash \Pi$, there are sets $\Delta \supseteq \Sigma$ and $\Gamma \supseteq \Pi$, such that $\Delta \nvdash \Gamma$, and Δ is a prime theory. To prove this, we enumerate the formulas of the language: $\beta_0, \beta_1, \beta_2, ...$, and define a sequence of sets Σ_n, Π_n $(n \in \omega)$ as follows. $\Sigma_0 = \Sigma$; $\Pi_0 = \Pi$. If $\Sigma_n \cup \{\beta_n\} \nvdash \Pi_n$, then $\Sigma_{n+1} = \Sigma_n \cup \{\beta_n\}$ and $\Pi_{n+1} = \Pi_n$. Otherwise $\Sigma_{n+1} = \Sigma_n$ and $\Pi_{n+1} = \Pi_n \cup \{\beta_n\}$. $\Delta = \bigcup_{n \in \omega} \Sigma_n$; $\Gamma = \bigcup_{n \in \omega} \Pi_n$.

It is not difficult to check by induction that for all n, $\Sigma_n \nvdash \Pi_n$. (Suppose this holds for n; if $\Sigma_n \cup \{\beta_n\} \nvdash \Pi_n$, the result for $n+1$ is immediate. So suppose that $\Sigma_n \cup \{\beta_n\} \vdash \Pi_n$ and $\Sigma_{n+1} \vdash \Pi_{n+1}$. Then $\Sigma_n \vdash \{\beta_n\} \cup \Pi_n$. By a sequence of moves that amount to "cut", $\Sigma_n \vdash \Pi_n$, contrary to induction hypothesis.) By compactness, it follows that $\Delta \nvdash \Gamma$.

It is also easy to check that Δ is a prime theory. Suppose that $\Delta \vdash \alpha$, but $\alpha \notin \Delta$. Then for some n, $\Sigma_n \cup \{\alpha\} \vdash \Pi_n$. Hence, $\Delta \vdash \Gamma$. Next, suppose that $\alpha \vee \beta \in \Delta$, but $\alpha \notin \Delta$ and $\beta \notin \Delta$. Then for some m and n, $\Sigma_n \cup \{\alpha\} \vdash \Pi_n$ and $\Sigma_m \cup \{\beta\} \vdash \Pi_m$. Hence $\Delta \cup \{\alpha \vee \beta\} \vdash \Gamma$, and so $\Delta \vdash \Gamma$.

4.4 Non-truth-functionality

Let us now return to the other basic approaches to paraconsistent logics. On the first of these, explosion is invalidated by employing a non-truth-functional account of negation. Typically, this account of negation is imposed on top of an orthodox account of positive logic. Thus, let an interpretation be a map, ν, from the set of formulas to $\{1, 0\}$, satisfying just the following conditions:

$\nu(\alpha \wedge \beta) = 1$ iff $\nu(\alpha) = 1$ and $\nu(\beta) = 1$
$\nu(\alpha \vee \beta) = 1$ iff $\nu(\alpha) = 1$ or $\nu(\beta) = 1$

In particular, the truth value of $\neg\alpha$ is independent of that of α. Validity is defined as truth preservation over all interpretations. It is obvious that explosion fails, since we may choose an evaluation that assigns both p and $\neg p$ (and their conjunction) the value 1, whilst assigning q the value 0.

These semantics can be characterised very simply in natural deduction terms by just the rules $\vee I$, $\vee E$, $\wedge I$ and $\wedge E$. Soundness is easy to check. For completeness, suppose that $\Sigma \nvdash \alpha$. Then put $\Pi = \{\alpha\}$, and extend Σ

to a prime theory, Δ, with the same property, as in 4.3. Define a map, ν, as follows:

$$\nu(\alpha) = 1 \text{ if } \alpha \in \Delta$$
$$\nu(\alpha) = 0 \text{ if } \alpha \notin \Delta$$

It is easy to check that ν is an interpretation. Hence, we have the result.

This system contains no inferences that involve negation essentially. For this reason, $\neg$ can hardly be thought of as a negation functor. Stronger paraconsistent systems, where this is more plausibly the case, can be obtained by adding conditions on the semantics. The following are some examples:[39]

(i) if $\nu(\alpha) = 0$, $\nu(\neg\alpha) = 1$
(ii) if $\nu(\neg\neg\alpha) = 1$, $\nu(\alpha) = 1$
(iii) if $\nu(\alpha) = 1$, $\nu(\neg\neg\alpha) = 1$
(iv) $\nu(\neg(\alpha \wedge \beta)) = \nu(\neg\alpha \vee \neg\beta)$
(v) $\nu(\neg(\alpha \vee \beta)) = \nu(\neg\alpha \wedge \neg\beta)$

Sound and complete rule systems can be obtained by adding the corresponding rules, which are, respectively:

(i) $\dfrac{}{\alpha \vee \neg\alpha}$

(ii) $\dfrac{\neg\neg\alpha}{\alpha}$

(iii) $\dfrac{\alpha}{\neg\neg\alpha}$

(iv) $\dfrac{\neg(\alpha \wedge \beta)}{\overline{\neg\alpha \vee \neg\beta}}$

(v) $\dfrac{\neg(\alpha \vee \beta)}{\overline{\neg\alpha \wedge \neg\beta}}$

(Double underlining indicates a two-way rule of inference, and a zero premise rule, as in (i), can be thought of as an assumption that discharges itself.) The corresponding soundness and completeness proofs are simple extensions of the basic arguments.

These additions give the $\wedge, \vee, \neg$-fragments of various systems in the literature. (i) gives that of Batens' *PI* [1980]; (i) and (ii) that of da Costa's C_ω;[40] (i)-(v) that of Batens' PI^S. In PI^S every sentence is logically equivalent to one in Conjunctive Normal Form. This can be used to show that PI^S

[39] Some others can be found in Loparić and da Costa [1984], and Beziau [1990].

[40] Semantics of the present kind for the da Costa systems were first proposed in da Costa and Alves [1977].

is a maximal paraconsistent logic, in the sense that any logic that extends it is not paraconsistent. (For details, see Batens [1980].)

Observe, for future reference, that if we add to PI or an extension thereof the condition:

if $\nu(\alpha) = 1$, $\nu(\neg\alpha) = 0$

then all interpretations are classical, and so we have classical logic. As may easily be checked, adding this is sound and complete with respect to the rule of inference:

$$\frac{\alpha \wedge \neg\alpha}{\beta}$$

Another major da Costa system, C_1, extends C_ω in accordance with the following idea. It should be possible to express in the language the idea that a sentence, α, behaves consistently; and for consistent sentences classical logic should apply. Let us write $\neg(\alpha \wedge \neg\alpha)$ as α^0. Then it is natural enough to suppose that α^0 expresses the consistency of α. It does not, in any of the above systems, since we may have $\alpha \wedge \neg\alpha \wedge \alpha^0$ true in an interpretation. This is exactly what is ruled out by the condition:

(vi) $\nu(\alpha) = \nu(\neg\alpha)$ then $\nu(\alpha^0) = 0$[41]

($\nu(\alpha) = \nu(\neg\alpha)$ iff both are 1, by semantic condition (i). Note that (i) also guarantees the converse of (vi): $\nu(\alpha) \neq \nu(\neg\alpha)$ then $\nu(\alpha^0) = 1$.)

C_1 is obtained by adding (vi) to C_ω, together with the following condition, which requires consistency to be preserved under syntactic constructions:

(vii) if $\nu(\alpha^0) = \nu(\beta^0) = 1$ then $\nu((\neg\alpha)^0) = \nu((\alpha \wedge \beta)^0) = \nu((\alpha \vee \beta)^0) = 1$

The deduction rules that correspond to (vi) and (vii) are, respectively:

$$\text{(vi)} \quad \frac{\alpha \wedge \neg\alpha \wedge \alpha^0}{\beta}$$

$$\text{(vii)} \quad \frac{\alpha^0}{(\neg\alpha)^0} \qquad \frac{\alpha^0 \quad \beta^0}{(\alpha \wedge \beta)^0} \qquad \frac{\alpha^0 \quad \beta^0}{(\alpha \vee \beta)^0}$$

Soundness and completeness of the extensions are easily checked.

Now suppose that we have a piece of valid classical reasoning concerning formulas composed of parameters $p_1, ..., p_n$. If we *assume* $p_1^0, ..., p_n^0$ then for

[41]Da Costa's actual condition is: if $\nu(\alpha^0) = \nu(\beta \supset (\alpha \wedge \neg\alpha)) = 1$ then $\nu(\beta) = 0$. This is equivalent, given his account of the conditional.

every such formula, α, α^0 follows by the appropriate applications of the rules of (vii). Hence, whenever we have established $\alpha \wedge \neg\alpha$ we may apply rule (vi) to give β. But the addition of this inference is sufficient to give classical logic, as I have already observed. Hence any valid classical reasoning may be recaptured formally by adding the appropriate consistency assumptions.[42]

One final comment on treating negation non-truth-functionally. It is a consequence of this that the substitutivity of provable equivalents breaks down in general. For example, even though α is logically equivalent to $\alpha \wedge \alpha$ there is no guarantee that the negations of these formulas have the same truth value in an interpretation.[43]

4.5 Many Values

The previous approach sticks with the traditional two truth values, and obtains a paraconsistent logic by making negation non-truth-functional. The next approach retains truth functionality, but drops the idea that there are exactly two truth values. That is, such logics are many-valued.[44] A many-valued logic will be paraconsistent if it is possible for a formula and its negation both to take designated values (whilst not everything does). A natural way of obtaining this is to have a designated value that is a fixed point for negation. The simplest such logic is a three valued one with values, t, b, and f, where t and b are designated, and the matrices are:

$\neg$	
t	f
b	b
f	t

$\wedge$	t	b	f
t	t	b	f
b	b	b	f
f	f	f	f

$\vee$	t	b	f
t	t	t	t
b	t	b	b
f	t	b	f

It will be noted that these are just the matrices of Łukasiewicz and Kleene's 3-valued logics, where the middle value is normally thought of as *undecidable*, or *neither true nor false*, and so not designated. It was the thought that this value might be read as *both true and false*—a natural enough thought, given dialetheism—and so be designated, that marks the start of many-valued paraconsistent logic. This was the approach proposed by Asenjo (see his [1966]), and others, e.g. Priest [1979], where the logic is called LP, a nomenclature that I will stick with here.

[42]It might be suggested that one ought not to take α^0 as expressing consistency unless it, itself, behaves consistently. This thought motivates the weaker da Costa system C_2, which is the same as C_1, except that α^0 is replaced everywhere by $\alpha^0 \wedge \alpha^{00}$. Of course, there is no reason to suppose that this expresses the consistency of α unless it, itself, behaves consistently. This thought motivates the da Costa system C_3 where α^0 is replaced everywhere by $\alpha^0 \wedge \alpha^{00} \wedge \alpha^{000}$. And so on for all the da Costa Systems C_i, for finite non-zero i.

[43]For a discussion of this in the context of da Costa's logics, see Urbas [1989].

[44]For a general discussion of many-valued logics, see the articles on the topic in this *Handbook*. See also, Rescher [1969].

LP may be generalised in various different ways. One is as follows. If we let $t = +1$, $b = 0$ and $f = -1$ then the truth conditions of LP are:

$$\nu(\neg\alpha) = -\nu(\alpha)$$
$$\nu(\alpha \wedge \beta) = \min\{\nu(a), \nu(\beta)\}$$
$$\nu(\alpha \vee \beta) = \max\{\nu(a), \nu(\beta)\}$$

The same conditions can be used for any set of integers, X, containing 0 and closed under $-$. The designated values are the non-negative values. Let us call this a *Sugihara generalisation*, after the person who, in effect, first proposed a matrix of this kind, where X was the set of all integers.[45]

Any Sugihara generalisation, though semantically different from LP, is essentially equivalent to it. Any LP countermodel is a Sugihara countermodel. But conversely, if we have a Sugihara countermodel, we can obtain an LP countermodel by mapping all positive values to $+1$, 0 to 0, and all negative values to -1. A little thought is sufficient to establish that the mapping respects the matrices and preserves designated values, as required.

A different way of generalising LP is as follows. If we let $t = 1$, $b = 0.5$ and $f = 0$ then the truth conditions of LP are:

$$\nu(\neg\alpha) = 1 - \nu(\alpha)$$
$$\nu(\alpha \wedge \beta) = \min\{\nu(a), \nu(\beta)\}$$
$$\nu(\alpha \vee \beta) = \max\{\nu(a), \nu(\beta)\}$$

The same conditions can be used for any set of reals $\{0, 0.5, 1\} \subseteq X \subseteq [0, 1]$, which is closed under subtraction (of a greater by a lesser). For suitable choices of X, these are the matrices of the odd-numbered finite Łukasiewicz many-valued logics, and for $X = [0, 1]$ they are the matrices of Łukasiewicz' continuum-valued logic. In Łukasiewicz' logics proper, the only designated value is 1, which does not give a paraconsistent logic. But if one takes the designated values to be $\{x; a < x \leq 1\}$ (or $\{x; a \leq x \leq 1\}$) then the logic will be paraconsistent provided that $0 < a < 0.5$ (or $0 < a \leq 0.5$). Let us call such logics *Łukasiewicz generalisations.* In a Łukasiewicz generalisation where the set of truth values is $[0, 1]$, these may naturally be thought of as degrees of truth. Hence, such a logic is a natural candidate for a paraconsistent fuzzy logic (logic of vagueness).[46]

It is not difficult to see that any Łukasiewicz generalisation is, in fact, equivalent to LP. As with the Sugihara generalisations, any LP countermodel is a Łukasiewicz countermodel; and conversely, any Łukasiewicz

[45] See Anderson and Belnap [1975], sect. 26.9.

[46] A variation on this theme is given by Peña in a number of papers. (See, e.g., Peña [1984].) Peña takes truth values to be an ordered set of more complex entities defined in terms of the interval $[0, 1]$.

countermodel can be collapsed into an LP countermodel by the mapping, f, defined thus:

$$\begin{aligned} f(x) &= 1 && \text{if } 1-a \leq x \leq 1 \\ &= 0.5 && \text{if } a < x < 1-a \\ &= 0 && \text{if } 0 \leq x \leq a \end{aligned}$$

or for the case where a is a designated value:

$$\begin{aligned} f(x) &= 1 && \text{if } 1-a < x \leq 1 \\ &= 0.5 && \text{if } a \leq x \leq 1-a \\ &= 0 && \text{if } 0 \leq x < a \end{aligned}$$

The generalisations of LP that we have considered in this section all, therefore, generate the same logic. What its proof theory is, we will see in the next.

4.6 Relational Valuations

Standardly, semantic evaluations are thought of as functions from formulas to truth values, say, 0 and 1. Another way of invalidating explosion is to take them to be, not functions, but relations. A formula may then relate to both 0 and 1, another way of expressing the thought that a sentence is both true and false. Assuming that negation behaves as usual, this means that both p and $\neg p$ may relate to 1, whilst an arbitrary formula may not. A natural way of spelling out this idea is as follows.

If P is the set of propositional parameters, an evaluation, ρ, is a subset of $P \times \{0,1\}$. The evaluation is extended to a relation for all formulas by the familiar looking recursive clauses:

$\neg\alpha\rho 1$ iff $\alpha\rho 0$
$\neg\alpha\rho 0$ iff $\alpha\rho 1$

$\alpha \wedge \beta\rho 1$ iff $\alpha\rho 1$ and $\beta\rho 1$
$\alpha \wedge \beta\rho 0$ iff $\alpha\rho 0$ or $\beta\rho 0$

$\alpha \vee \beta\rho 1$ iff $\alpha\rho 1$ or $\beta\rho 1$
$\alpha \vee \beta\rho 0$ iff $\alpha\rho 0$ and $\beta\rho 0$

Let us say that a formula, α, is true in an interpretation, ρ, iff $\alpha\rho 1$, and false iff $\alpha\rho 0$; then validity may be defined as truth preservation in all interpretations. According to this account, classical logic is just the special case where multi-valued relations have been forgotten.

These semantics are the Dunn semantics for the logic of First Degree Entailment, FDE.[47] In natural deduction terms, FDE can be characterised by the rules $\wedge I$, $\wedge E$, $\vee I$ and $\vee E$, together with the rules:

$$\frac{\neg(\alpha \wedge \beta)}{\overline{\neg\alpha \vee \neg\beta}} \qquad \frac{\neg\alpha \wedge \neg\beta}{\overline{\neg(\alpha \vee \beta)}} \qquad \frac{\alpha}{\overline{\neg\neg\alpha}}$$

Soundness is easily checked. For completeness, suppose that $\Sigma \not\vdash \alpha$. Extend Σ to a prime theory, Δ, with the same property, as in 4.3. Now define an interpretation, ρ, thus:

$p\rho 1$ iff $p \in \Delta$
$p\rho 0$ iff $\neg p \in \Delta$

A straightforward (joint) induction shows that this characterisation extends to all formulas. Completeness follows.

There are two natural restrictions that one may place upon Dunn evaluations:

#1 for every p, there is at most one x such that $p\rho x$
#2 for every p there is at least one x such that $p\rho x$

Both conditions extend from propositional parameters to all formulas, by a simple induction. Thus, the first condition ensures that the relation is *functional*; the second that it is *total*. A relation that satisfies both conditions is just a classical evaluation.

These extra conditions are sound and complete with respect to the extra rules:

$$\frac{\alpha \wedge \neg\alpha}{\beta} \qquad \frac{}{\alpha \vee \neg\alpha}$$

respectively, as simple extensions of the completeness proofs demonstrate.

We can express the relational semantics in functional terms by taking an evaluation to be a function from formulas to *subsets* of $\{1, 0\}$, since there is an obvious isomorphism between relations, ρ, and functions, ν, given by the condition:

$\alpha\rho x$ iff $x \in \nu(\alpha)$

[47]Published in Dunn [1976], though he discovered them somewhat earlier than this. In the present context, it might be better to call the system 'Zero Degree Entailment' since the language does not contain a conditional connective.

In this way, FDE can be seen as a many- (in fact, four-) valued logic.[48]

Restriction #2, which ensures that no formula takes the value ϕ, gives a three-valued logic that is identical with LP. It is easy enough to check that the values $\{1\}$, $\{1,0\}$, and $\{0\}$ work the same way as t, b, and f, respectively. I will make this identification in the rest of this essay. Restriction #1, which ensures that no formula takes the value $\{1,0\}$, obviously gives an explosive logic, which is, in fact, the strong Kleene three-valued logic. This is therefore a logic dual to LP.[49]

A feature of these semantics for LP and FDE is that they are monotonic in the following sense. Let ν_1 and ν_2 be functional evaluations. If for all propositional parameters, p, $\nu_1(p) \subseteq \nu_2(p)$ then for all α, $\nu_1(\alpha) \subseteq \nu_2(\alpha)$. The proof of this is by a simple induction. One consequence of this for LP is worth remarking on. LP is clearly a sub-logic of classical logic, since it has the classical matrices as sub-matrices. The consequence relation of LP is weaker than that of classical logic, since it is paraconsistent. But the set of logical truths of LP is identical with that of classical logic. For suppose that α is not valid in LP. Let ν be an evaluation such that $1 \notin \nu(\alpha)$. Let ν' be the interpretation that is the same as ν, except that for every parameter, p, if $\nu(p) = \{0,1\}$, $\nu'(p) = \{0\}$. This is a classical evaluation; and by monotonicity, $1 \notin \nu'(\alpha)$, as required.

Another feature of these semantics is the evaluation that assigns every propositional parameter the value $\{1,0\}$, $v_{\{1,0\}}$; and, in the four-valued case, the evaluation that gives every parameter the value ϕ, v_ϕ. A simple induction shows that these properties extend to all formulas. Thus, $v_{\{1,0\}}$ makes all formulas true—and false—and v_ϕ makes every formula neither. In particular, then, FDE has no logical truths.[50]

4.7 Possible Worlds

Yet another, closely connected, way of invalidating explosion is to treat negation as an intensional operator. This way was proposed by the Routleys in [1972]. A *Routley interpretation* is a structure, $\langle W, *, \nu\rangle$, where W is a set (of worlds), $*$ is a map from W to W, and ν maps sets of pairs comprising a world and propositional parameter to $\{1,0\}$. (I will write $\nu(w,\alpha)$ as $\nu_w(\alpha)$.) The truth conditions for conjunction and disjunction are the standard:

$$\nu_w(\alpha \wedge \beta) = 1 \text{ iff } \nu_w(\alpha) = 1 \text{ and } \nu_w(\beta) = 1$$

[48] In fact, the straight truth tables with values 1, 2, 3 and 4 were enunciated by Smiley. See Anderson and Belnap [1975], p. 161.

[49] I will usually use the functional semantic representation for FDE and LP in the rest of this essay. A word of warning, though: in the context of a dialetheic metatheory, the functional approach may have consequences that the relational approach, proper, does not have. See Priest and Smiley [1993], p. 49ff.

[50] Further interesting properties of LP and FDE are established in Pynko [1995a] and [1995b].

$$\nu_w(\alpha \vee \beta) = 1 \text{ iff } \nu_w(\alpha) = 1 \text{ or } \nu_w(\beta) = 1$$

The truth conditions for negation are:

$$\nu_w(\neg\alpha) = 1 \text{ iff } \nu_{w^*}(\alpha) = 0$$

Note that if $w^* = w$, these conditions just reduce to the classical ones. A natural understanding of the $*$ operator is a moot point.[51] I will return to the issue in a moment. Validity is defined in terms of truth preservation at all worlds of all interpretations.

In natural deduction terms, this system can be characterised by modifying that for FDE by dropping the rule for double negation, and replacing it with:

$$\dfrac{\begin{array}{c}\overline{\alpha}\\ \vdots \\ \beta\end{array} \qquad \neg\beta}{\neg\alpha}$$

where, in the subproof, there are no undischarged assumptions other than α. Soundness is easily checked. For completeness, suppose that $\Sigma \not\vdash \alpha$. Extend Σ to a prime theory, Δ, with the same property, as in 4.3. Now define an interpretation, $\langle W, *, \nu\rangle$, where W is the set of all prime theories, $*$ is defined by the condition:

$$\alpha \in \Delta^* \text{ iff } \neg\alpha \notin \Delta$$

and ν is defined by:

$$\nu_\Delta(p) = 1 \text{ iff } p \in \Delta \qquad (\#)$$

It is not difficult to check that if Δ is a prime theory, so is Δ^* and hence that $*$ is well defined. First, suppose that $\alpha \notin \Delta^*$ and $\beta \notin \Delta^*$. Then $\neg\alpha$ and $\neg\beta$ are in Δ. Since Δ is a theory, $\neg\alpha \wedge \neg\beta \in \Delta$, and so $\neg(\alpha \vee \beta) \in \Delta$. Hence, $\alpha \vee \beta \notin \Delta^*$. Next, suppose that $\Delta^* \vdash \alpha$, but $\alpha \notin \Delta^*$. Then for some $\beta_1, ..., \beta_n \in \Delta^*$, $\beta_1 \wedge ... \wedge \beta_n \vdash \alpha$. Hence, by contraposition and De Morgan $\neg\alpha \vdash \neg\beta_1 \vee ... \vee \neg\beta_n$. But $\neg\alpha \in \Delta$; hence $\neg\beta_1 \vee ... \vee \neg\beta_n \in \Delta$. Since Δ is prime, for some $1 \leq i \leq n$, $\neg\beta_i \in \Delta$, i.e., $\beta_i \notin \Delta^*$. Contradiction.

An easy recursion shows that $(\#)$ extends to all formulas. The result follows.

[51] For some discussion and references, see the article on Relevance Logic and Entailment in this *Handbook*.

The logic can be made stronger without (necessarily) ruining its paraconsistency by adding further conditions on $*$. The most notable is: $w^{**} = w$. This is sound and complete with respect to the additional rule:

$$\frac{\alpha}{\overline{\neg\neg\alpha}}$$

as a simple extension of the completeness argument demonstrates.

These semantics are, in fact, very closely related to the those for FDE of the previous section. Given an FDE interpretation, ν, define a Routley evaluation on the worlds w and w^*, as follows:

$\nu_w(p) = 1$ iff $1 \in \nu(p)$
$\nu_{w^*}(p) = 1$ iff $0 \notin \nu(p)$

A simple induction shows that these conditions follow for all formulas. Conversely, we can turn the conditions into reverse. Given any Routley evaluation on a pair of worlds, w, w^*, define a Dunn evaluation by the conditions:

$1 \in \nu(p)$ iff $\nu_w(p) = 1$
$0 \in \nu(p)$ iff $\nu_{w^*}(p) \neq 1$

Essentially the same induction shows that these conditions hold for all formulas. Hence, the two semantics are inter-translatable, and validate the same proof theories.[52] The translation also suggests a natural interpretation of the $*$ operator. w^* is that world characterised by the set of unfalsehoods of w. (This is, of course, in general, distinct from the set of truths in a four-valued context.)

Under the above translation, the condition: $1 \in \nu(p)$ or $0 \in \nu(p)$, which gives an LP interpretation, is equivalent to: $\nu_w(p) = 1$ or $\nu_{w^*}(p) \neq 1$; imposing which condition on an intensional interpretation therefore gives an intensional semantics for LP.

4.8 Algebraic Semantics

Let us now turn to the final approach to paraconsistent logics that we will consider, an algebraic one. In algebraic logic, an interpretation is a homomorphism, ν, from sentences into some algebraic structure, $\mathcal{A} = \langle A, \wedge, \vee, \neg \rangle$; i.e., $\nu(\neg\alpha) = \neg\nu(\alpha)$, $\nu(\alpha \wedge \beta) = \nu(\alpha) \wedge \nu(\beta)$, etc. (I will use the same signs for the connectives and the algebraic operations. Context, and the style of variable, will serve to disambiguate.) If the algebra is a lattice—as it usually

[52]Which shows that the contraposition rule is admissible in FDE, something that is not at all obvious.

is, and will be in all the cases we consider—the consequence relation of the logic is represented by the lattice order relation, defined in the usual way: $a \leq b$ iff $a \wedge b = a$. Thus, a logic will be paraconsistent if it is possible in the algebra to have an a and b such that $a \wedge \neg a \not\leq b$.

Several of the logics that we have looked at can be algebraicised. Consider, for example, FDE. If we take the four-valued semantics for this, we can think of the values as a lattice whose Hasse diagram is as follows:

$$\begin{tikzcd}
& \{1\} & \\
\{1,0\} \arrow[ur] & & \phi \arrow[ul] \\
& \{0\} \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

($\wedge$ is lattice-meet; $\vee$ is lattice join; $\neg\{1,0\} = \{1,0\}$, and $\neg\phi = \phi$.) This generalises to a *De Morgan algebra.* A De Morgan algebra is a structure $\mathcal{A} = \langle A, \wedge, \vee \rangle$, where $\langle A, \wedge, \vee, \neg \rangle$ is a distributive lattice, and $\neg$ is an involution of period 2, i.e.:

$$\neg\neg a = a$$
$$a \leq b \Rightarrow \neg b \leq \neg a$$

The structures take their name from the fact that in every such algebra $\neg(a \wedge b) = \neg a \vee \neg b$ holds, as do the other De Morgan laws.

Define an inference $\alpha_1, ..., \alpha_n/\beta$ to be algebraically valid iff for every homomorphism, ν, into a De Morgan algebra, $\mathcal{A}$, $\nu(\alpha_1) \wedge ... \wedge \nu(\alpha_n) \leq \nu(\beta)$. Then the algebraically valid inferences are exactly those of FDE. It is easy to check that the rule system for FDE is sound with respect to these semantics. Completeness follows from completeness in the four-valued case. Alternatively, we can give a direct argument as follows.

Consider the relation $\alpha \sim \beta$, defined by: $\alpha \vdash \beta$ and $\beta \vdash \alpha$. One can check that this is an equivalence relation, and a congruence on the logical operators (i.e., if $\alpha_1 \sim \beta_1$ and $\alpha_2 \sim \beta_2$ then $\alpha_1 \wedge \alpha_2 \sim \beta_1 \wedge \beta_2$, etc.).[53] If F is the set of formulas, define the quotient algebra, $\mathcal{A} = \langle F/\sim, \wedge, \vee, \neg \rangle$, where, if $[\alpha]$ is the equivalence class of α, $\neg[\alpha] = [\neg\alpha]$, $[\alpha] \wedge [\beta] = [\alpha \wedge \beta]$, etc. One can check that $\mathcal{A}$ is a De Morgan lattice. Now, let ν be the homomorphism that maps every α to $[\alpha]$. If $\nu(\alpha_1) \wedge ... \wedge \nu(\alpha_n) \leq \nu(\beta)$. Then $[\alpha_1 \wedge ... \wedge \alpha_n] \leq [\beta]$, i.e., $[\alpha_1 \wedge ... \wedge \alpha_n \wedge \beta] = [\alpha_1 \wedge ... \wedge \alpha_n]$. Hence, $\alpha_1 \wedge ... \wedge \alpha_n \vdash \alpha_1 \wedge ... \wedge \alpha_n \wedge \beta$ and so $\alpha_1 \wedge ... \wedge \alpha_n \vdash \beta$. Conversely, then, if $\alpha_1 \wedge ... \wedge \alpha_n \nvdash \beta$ then $\nu(\alpha_1) \wedge ... \wedge \nu(\alpha_n) \not\leq \nu(\beta)$, as required.

[53]The only tricky point concerns negation. For this, we need to appeal to the fact, which we have already noted, that if $\alpha \vdash \beta$ then $\neg\beta \vdash \neg\alpha$. This can be established directly, by an induction on proofs.

It should be noted that not all the logics we have considered in previous sections algebraicise. In particular, the non-truth-functional logics resist this treatment in general. This is for the same reason that the substitutivity of provable equivalents breaks down: the semantic value of $\neg\alpha$ is *entirely independent* of that of α. It cannot, therefore, correspond to any well-defined algebraic operation.

The point can be made more precise in many cases. Suppose that $\mathcal{A}$ is some algebraic structure for a logic, and consider any interpretation, ν, with values in the algebra, such that for some p, q and r, $\nu(p) = \nu(q) \neq \nu(r)$. Then the condition $\nu(\alpha) = \nu(\beta)$ is a congruence relation on the set of formulas, and collapse by it gives a non-degenerate quotient algebra (i.e., an algebra that is neither a single-element algebra, nor the algebra of formulas). But many non-truth-functional logics can be shown to have no such thing. (See, e.g., Mortensen [1980].)

One final algebraic paraconsistent logic is worth noting. This is that of Goodman [1981]. A Heyting algebra can be thought of as a distributive lattice, with a bottom element, $\bot$, and an operator, $\rightarrow$, satisfying the condition:

$$a \wedge b \leq c \text{ iff } a \leq b \rightarrow c$$

(which makes $\bot \rightarrow \bot$ the top element). We may define $\neg a$ as $a \rightarrow \bot$.

Let $\mathcal{T}$ be a topological space. Then a standard example of a Heyting algebra is the topological Heyting algebra $\langle X, \wedge, \vee, \rightarrow, \bot\rangle$, where X is the set of open sets in $\mathcal{T}$, $\wedge$ and $\vee$ are intersection and union, respectively, $\bot$ is ϕ, and $a \rightarrow b$ is $(\overline{a} \vee b)^o$—overlining denotes complementation and ^{o} is the interior operator of the topology. $\neg a$ is clearly $\overline{a}^o$.

It is well known that for finite sets of premises, Intuitionistic logic is sound and complete with respect to the class of Heyting algebras, in fact, with respect to the topological Heyting algebras. That is, $\alpha_1, ... \alpha_n \models_I \beta$ iff for every homomorphism, ν, into such an algebra, $\nu(\alpha_1 \wedge ... \wedge \alpha_n) \leq \nu(\beta)$.[54]

The whole construction can be dualised in a natural way to give a paraconsistent logic. A dual Heyting algebra is a distributive lattice, with a top element, $\top$, and an operator, $\leftarrow$, satisfying the condition:

$$a \leq b \vee c \text{ iff } a \leftarrow b \leq c$$

(which makes $\top \leftarrow \top$ the bottom element). We may define $\neg a$ as $\top \leftarrow a$. As may be checked, if $\mathcal{T}$ is a topological space, then the structure $\langle X, \wedge, \vee, \leftarrow, \top\rangle$ is a dual Heyting algebra, where X is the set of *closed* sets of $\mathcal{T}$, $\wedge$ and $\vee$ are intersection and union, respectively, $\top$ is the whole space,

[54] See, e.g., Dummett [1977], 5.3.

and $a \leftarrow b$ is $(a \wedge \bar{b})^c$—where ^{c} is the closure operator of the topology. $\neg b$ is clearly $\bar{b}^c$.

The logic generated by dual Heyting algebras is dual to Intuitionistic logic. In particular, in Intuitionistic logic we have $\alpha \wedge \neg\alpha \models \beta$, but not $\alpha \models \beta \vee \neg\beta$; and $\alpha \models \neg\neg\alpha$, but not $\neg\neg\alpha \models \alpha$. Thus in dual Intuitionist logic, we have $\alpha \models \beta \vee \neg\beta$ but not $\alpha \wedge \neg\alpha \models \beta$; and $\neg\neg\alpha \models \alpha$ but not $\alpha \models \neg\neg\alpha$. For a topological counter-model to the first, consider the real line with its usual topology, and an interpretation, ν, that maps p to $[-1, +1]$, and q to ϕ. Then $\nu(p \wedge \neg p) = \{-1, +1\} \not\subseteq \phi = \nu(q)$. (This illustrates how the points in the set represented by $p \wedge \neg p$ may be thought of as the points on the topological boundary between the set of points represented by p and the set of points represented by $\neg p$.) For a counter-model to the second, let $\nu(p) = \{0\}$. Then $\nu(\neg\neg p) = \phi \not\supseteq \nu(p)$.

If $\models \alpha$ in dual Intuitionist logic, then $\models_C \alpha$, since the two-element Boolean algebra is a dual Heyting algebra. Conversely, if α is any classical tautology, its dual, α', is a contradiction. Hence, $\models_C \neg\alpha'$. But then by a result of Glivenko, $\models_I \neg\alpha'$, and so $\alpha' \models_I$. Thus by duality, in dual Intuitionist logic $\models \alpha$. The logical truths of dual Intuitionist logic are therefore the same as those of classical logic.

It is worth noting that just as Intuitionist logic can be given an intensional semantics, namely Kripke semantics, so can dual Intuitionist logic; we simply dualise the Kripke construction. For further details of all the above, see Goodman [1981].[55]

5 CONDITIONAL CONNECTIVES

We have now looked at most of the basic techniques of paraconsistent logic, applied to languages containing only negation, conjunction and disjunction.[56] I will call this language the *basic* language. Next, we will look at some important extensions of these techniques (which do not ruin paraconsistency). In this part, we will start with the conditional, by which I mean some con-

[55] It is well known that in a certain well defined sense, Intuitionist logic can be seen as the "internal logic" of the category-theoretic structures called topoi. It is possible to dualise the construction involved there to show that dual Intuitionist logic has an equally good claim to that title. For details, see Mortensen [1995], who calls the $\wedge$, $\vee$, $\neg$-fragment of a dual Heyting algebra a 'paraconsistent algebra'.

[56] There are others, such as the use of the techniques of combinatorial logic, but I will not go into these here. For details, one can consult, e.g., Bunder [1984]. There *ought* to be yet more. The discussion of connexivism in 3.2 suggests that there ought to be a distinctive connexivist approach to paraconsistency. To date, this has not emerged. The most articulated modern connexivist logic is due to McCall (see sect. 29.8 of Anderson and Belnap [1975], which can also be consulted for references to other discussions). Although this provides a connexivist treatment of the connective $\rightarrow$, the logic of the basic language is classical, and so explosive. Alternatively, one can formulate versions of relevant logic that contain connexivist principles. See Routley [1978] and Mortensen [1984].

nective, $\rightarrow$ (if necessary, added to the basic language), satisfying, at least, *modus ponens*: $\alpha, \alpha \rightarrow \beta \models \beta$.

Although paraconsistency does not concern the conditional as such, many of the paraconsistent logics that we have looked at have distinctive approaches to the conditional. And this is no accident. If one identifies $\alpha \rightarrow \beta$ with the material conditional, $\alpha \supset \beta$, defined in the usual way as $\neg\alpha \vee \beta$, then *modus ponens* reduces to the disjunctive syllogism. But in any logic where disjunction behaves normally and deducibility is transitive, the disjunctive syllogism must fail, or explosion would arise, due to the "Lewis independent argument". Specifically, in all the logics we have looked at except filter logics and some of the non-adjunctive logics, the syllogism fails. In such logics, therefore, a distinct account of the conditional is required. For completeness' sake, we will start by considering the others.

5.1 $\rightarrow$ *as* $\supset$

In filter logics, we may simply identify $\rightarrow$ with $\supset$. Things then proceed as before. A one-premise inference in this language, α/β, is prevalid iff it is classically valid and $F(\alpha, \beta)$. It is valid iff it is a substitution instance of a prevalid inference.[57]

In the natural extension of Tennant's semantic approach, an inference from Π to Σ is prevalid iff $\Pi \models_C \Sigma$ and for no proper subsets of Π and Σ, Π' and Σ', respectively, $\Pi' \models_C \Sigma'$. The natural extension of the proof theory is to add the conditional rules:

$$\frac{\Pi, \alpha : \beta, \Delta}{\Pi : \alpha \rightarrow \beta, \Delta} \qquad \frac{\Pi_1, \alpha : \Delta_1 \quad \beta, \Pi_2 : \Delta_2}{\Pi_1, \Pi_2, \alpha \rightarrow \beta : \Delta_1, \Delta_2}$$

Unfortunately, the equivalence between these two approaches now fails. For, semantically, $p \models \neg p \rightarrow q$ (though the system is still paraconsistent); but without dilution there is no proof of the sequent $p : \neg p \rightarrow q$. At this point, Tennant prefers to go with the proof theory rather than the semantics. He also prefers the intuitionist version, which allows at most one formula on the right-hand side of a sequent. For further details, including natural deduction versions of the proof theory, see Tennant [1987], ch. 23.

In [1992] Tennant suggests modifying the rule for the introduction of $\rightarrow$ on the right.[58] The α in the premise sequent is made optional, and the following rule is added.

[57]One can modify this approach, invoking the filter in the truth conditions of the conditional itself, to give logics of a more relevant variety. This is pursued in a number of the essays in *Philosophical Studies* **26** (1979), no. 2, a special issue on relatedness logics.

[58]In fact, he gives the natural deduction rules. The sequent rules described are the obvious equivalents.

$$\frac{\Pi, \alpha : \Delta}{\Pi : \alpha \rightarrow \beta, \Delta}$$

The exact relationship between these rules and the above semantics is as yet unresolved.

In the non-adjunctive logics of Rescher and Manor, and Schotch and Jennings: $\rightarrow$ may again be identified with $\supset$, producing no novelties. The machinery of maximally consistent subsets and partitions carries straight over.

5.2 Discursive Implication

The situation is otherwise with discursive logic. Here a distinct approach is required, since, as we have already seen, the disjunctive syllogism fails discursively.

Given an $S5$ interpretation, Jaśkowski adds a conditional, $\rightarrow$ (often written as $\supset_d$, and called discursive implication), and defines $\alpha \rightarrow \beta$ as $\Diamond\alpha \supset \beta$.[59] It is easy to check that in discursive logic $\alpha, \alpha \rightarrow \beta \models \beta$, since $\Diamond\alpha, \Diamond(\Diamond\alpha \supset \beta) \models_{S5} \Diamond\beta$ (and so there are essentially multi-premise inferences).

In fact, the logical truths of the pure $\rightarrow$ fragment of discursive logic are the same as those of the pure $\supset$ fragment of classical logic. For let $\alpha_\supset$ be any sentence containing only $\supset$s, and let $\alpha_\rightarrow$ be the corresponding sentence containing only $\rightarrow$s. In an $S5$ interpretation with only one world, $\alpha_\supset$ and $\Diamond\alpha_\rightarrow$ are equivalent. So if $\alpha_\supset$ is not a classical logical truth, $\Diamond\alpha_\rightarrow$ is not a discursive one. Conversely, suppose that $\alpha_\supset$ is a classical logical truth. We need to show that $\Diamond\alpha_\rightarrow$ is valid in every $S5$ model. As may easily be checked, in $S5$, $\Diamond(\Diamond\alpha \supset \beta)$ is logically equivalent to $\Diamond\alpha \supset \Diamond\beta$. Hence, given $\Diamond\alpha_\rightarrow$, we may "drive the $\Diamond$s inwards" to obtain a logically equivalent sentence where the modal operator applies only to propositional parameters. But this is a substitution instance of $\alpha_\supset$, and hence valid in $S5$. This result does not carry over to the full language. For example, $\not\models \alpha \rightarrow (\neg\alpha \rightarrow \beta)$, since, as may be checked, $\not\models_{S5} \Diamond(\Diamond\alpha \supset (\Diamond\neg\alpha \supset \beta))$.[60]

Full discursive logic can naturally be generalised in two obvious ways. The first is by using some modal logic other than $S5$. The second is by changing the definition of what it is for a sentence, α, to hold discursively in an interpretation. We change this from $\Diamond\alpha$ holding to $M\alpha$ holding, where M is some other modality (i.e., string of $\Diamond$s and $\Box$s). For references and discussion, see Błaszczuk [1984] and Kotas and da Costa [1989].

[59] Given what amounts to Jaśkowski's identification of truth with truth in some possible world, it might be more natural to define $\alpha \rightarrow \beta$ as $\Diamond\alpha \rightarrow \Diamond\beta$. This would have just the same consequences.

[60] The natural definition of the biconditional, $\alpha \leftrightarrow \beta$, is $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)$. For reasons not explained, Jaśkowski defines it as $(\alpha \rightarrow \beta) \wedge (\beta \rightarrow \Diamond\alpha)$. This asymmetric and counter-intuitive definition would seem to have no significant advantages.

5.3 Da Costa's C-systems

The natural way of extending the non-truth functional semantics of 4.4 to include a conditional connective, in keeping with the idea that such logics are just the addition of a non-truth-functional negation to a standard positive logic, is to give $\rightarrow$ the classical truth conditions:

$$\nu(\alpha \rightarrow \beta) = 1 \text{ iff } \nu(\alpha) = 0 \text{ or } \nu(\beta) = 1$$

(Note that $\rightarrow$, so defined, is distinct from $\supset$.) Adding this condition to the logics of 4.4 (except, C_ω, which we will come to in a moment) gives the full (propositional) versions of the logics mentioned there; in particular it gives the da Costa logic C_1 (and the other C_i for finite non-zero i). In each case, a natural deduction system can be obtained by adding the rules:

$$\rightarrow E \quad \frac{\alpha \qquad \alpha \rightarrow \beta}{\beta}$$

$$\text{(a)} \quad \frac{\beta}{\alpha \rightarrow \beta}$$

$$\text{(b)} \quad \overline{\alpha \vee (\alpha \rightarrow \beta)}$$

Soundness is proved as usual. The extension to the completeness proof amounts to checking that for a prime theory, Σ, $\alpha \rightarrow \beta \in \Sigma$ iff $\alpha \notin \Sigma$ or $\beta \in \Sigma$. From left to right, the result follows by ($\rightarrow E$). From right to left: if $\beta \in \Sigma$ then the result follows from (a); if $\alpha \notin \Sigma$ then $(\alpha \rightarrow \beta) \in \Sigma$ by (b) and primeness.

If instead of (a) and (b), we add to any of these systems—except the ones with a consistency operator; I will come to these in a second—the rule:

$$\rightarrow I \quad \begin{array}{c} \overline{\alpha} \\ \vdots \\ \beta \\ \hline \alpha \rightarrow \beta \end{array}$$

we obtain, not classical positive logic, but intuitionist positive logic. (These rules are well known to be complete with respect to this logic.) In particular, if we add $\rightarrow I$ and $\rightarrow E$ to the rule system for the basic language fragment of C_ω we obtain da Costa's C_ω.

The intuitionist conditional is not, of course, truth functional, but a valuational semantics for C_ω can be obtained as follows. A *semi-valuation* is any function that satisfies the conditions for conjunction, disjunction and negation, plus:

if $\nu(\alpha \to \beta) = 1$ then $\nu(\alpha) = 0$ or $\nu(\beta) = 1$
if $\nu(\alpha \to \beta) = 0$ then $\nu(\beta) = 0$

A valuation is any semi-valuation, ν, satisfying the following condition. Let α be of the form $\alpha_1 \to (\alpha_2 \to (\alpha_3 ... \to \alpha_n)...)$, where α_n is not itself of the form $\beta \to \gamma$. Then if $\nu(\alpha) = 0$ there is a semi-valuation, ν', such that for all $1 \leq i < n$, $\nu'(\alpha_i) = 1$, and $\nu'(\alpha_n) = 0$. C_ω is sound and complete with respect to this notion of valuation. For details, see Loparić [1986].[61]

Changing the deduction rules for $\to$ to the intuitionist ones, makes no difference for those logics that contain a consistency operator, and in particular, the da Costa logics C_i for finite i.[62] The reason, *in nuce*, is that the consistency operator allows us to define a negation with the properties of classical negation. As is well known, the addition of such a negation to positive intuitionist logic is not conservative, but produces classical logic. In more detail, the argument for C_1 is as follows.[63]

Define $\neg^*\alpha$ as $\neg\alpha \wedge \alpha^o$. Then it is easy to check that:

$$\nu(\neg^*\alpha) = 1 \text{ iff } \nu(\alpha) = 0$$

In particular, then, $\neg^*\alpha$ satisfies the rules for classical negation:

$$\frac{}{\alpha \vee \neg^*\alpha} \qquad\qquad \frac{\alpha \wedge \neg^*\alpha}{\beta}$$

Given these, it is easy to show that $\alpha \to \beta \dashv\vdash \neg^*\alpha \vee \beta$. (Hint: from left to right, assume $\alpha \vee \neg^*\alpha$ and argue by cases. From right to left, assume α and $\neg^*\alpha \vee \beta$, and argue to β by cases.) Hence, $\to$ has the classical truth conditions.

5.4 Many-valued Conditionals

There are numerous ways to define a many-valued conditional operator. We will just look at two of the more systematic.[64]

Given a Sugihara generalisation of LP, one can define a conditional with the following truth conditions:

[61] A Kripke-style semantics for C_ω can be found in Baaz [1986].
[62] This was first observed, in effect, by da Costa and Guillaume [1965].
[63] The argument for the other C_is is similar.
[64] In the three-valued case, other definitions give the system of Asenjo and Tamburino [1975], and the J systems of D'Ottaviano and da Costa [1970]. A natural many-valued conditional, given the four-valued semantics of FDE, produces the system $BN4$ of Brady [1982].

$$\begin{aligned} \nu(\alpha \to \beta) &= \nu(\neg\alpha \vee \beta) && \text{if } \nu(\alpha) \leq \nu(\beta) \\ &= \nu(\neg\alpha \wedge \beta) && \text{if } \nu(\alpha) > \nu(\beta) \end{aligned}$$

This definition gives rise to "semi-relevant" logics, i.e., logics that avoid the standard paradoxes of relevance, but are still not relevant.

In the case where the set of truth values is the set of all integers, this gives the Anderson/Belnap logic RM. Proof-theoretically, RM is obtained from the relevant logic R, which we will come to in the next section, by adding the "mingle" axiom:

$$\vdash \alpha \to (\alpha \to \alpha)$$

For details of proofs, see Anderson and Belnap [1975], sect. 29.3.

In the 3-valued case, where the set of truth values is $\{-1, 0, +1\}$, the conditions for $\to$ give the matrix:

$\to$	$+1$	0	-1
$+1$	$+1$	-1	-1
0	$+1$	0	-1
-1	$+1$	$+1$	$+1$

and the stronger logic called $RM3$. This is sound and complete with respect to the axiomatic system obtained by augmenting the system R with the axioms:

$$\vdash (\neg\alpha \wedge \beta) \to (\alpha \to \beta)$$
$$\vdash \alpha \vee (\alpha \to \beta)$$

For the proof, see Brady [1982].

Turning to the second systematic approach, consider any Łukasiewicz generalisation of LP. Łukasiewicz' truth conditions for his conditional, $\mapsto$, are as follows:

$$\begin{aligned} \nu(\alpha \mapsto \beta) &= 1 && \text{if } \nu(\alpha) \leq \nu(\beta) \\ &= 1 - (\nu(\alpha) - \nu(\beta)) && \text{if } \nu(\alpha) > \nu(\beta) \end{aligned}$$

In the three-valued case, this gives the well known matrix:

$\mapsto$	1	0.5	0
1	1	0.5	0
0.5	1	1	0.5
0	1	1	1

Now the most notable feature of the Łukasiewicz definition, given that 0.5 is designated, is that *modus ponens* fails. For example, consider a valuation, ν, where $\nu(p) = 0.5$ and $\nu(q) = 0$. Then $\nu(p \mapsto q) = 0.5$. Hence $p, p \mapsto q \not\models q$. (*Modus ponens* is valid provided that the only designated value is 1, but then the logic is not paraconsistent.)

Kotas and da Costa [1978] get around this problem by adding to the language a new operator, Δ, with the truth conditions:

$$\begin{aligned} \nu(\Delta\alpha) &= 1 \quad \text{if } \nu(\alpha) \text{ is designated} \\ &= 0 \quad \text{otherwise} \end{aligned}$$

and then define a conditional, $\alpha \to \beta$, as $\Delta\alpha \mapsto \beta$.[65] They point out the similarity of this definition to Jaśkowski's definition of discursive implication. (In fact, they use the symbol $\Diamond$ instead of Δ because of this.)[66]

It is not difficult to check that *modus ponens* for $\to$ holds. In fact, as Kotas and da Costa point out, the $\wedge$, $\vee$, $\to$-fragment of the logic is exactly positive classical logic. The easiest way to see this is just to collapse the designated values to 1, and the others to 0, to obtain classical truth tables.

5.5 *Relevant* $\to$*s*

Given a Routley interpretation (say one for *FDE*, though the other cases will be similar), it is natural to treat $\to$ intensionally. The simplest way is to give it the $S5$ truth conditions:

$$\nu_w(\alpha \to \beta) = 1 \text{ iff for all } w' \in W\ (\nu_{w'}(\alpha) = 1 \Rightarrow \nu_{w'}(\beta) = 1)$$

Clearly, given an interpretation either $\alpha \to \beta$ is true at all worlds, or at none. With the Routley $*$ giving the semantics for negation, it follows that the same is true of negated conditionals. It also follows that $\nu_w(\alpha \to \beta) = 1$ iff $\nu_{w^*}(\alpha \to \beta) = 1$ iff $\nu_w \neg(\alpha \to \beta) \neq 1$. Thus, the semantics validate the rules:

$$\text{LEM}_{\to} \qquad \frac{}{(\alpha \to \beta) \vee \neg(\alpha \to \beta)}$$

$$\text{EFQ}_{\to} \qquad \frac{\alpha \to \beta \qquad \neg(\alpha \to \beta)}{\gamma}$$

and so are unsuitable for serious paraconsistent purposes. Moreover, even though there may be worlds where $\alpha \wedge \neg\alpha$ is true, or where $\alpha \vee \neg\alpha$ is false,

[65] In fact, their treatment is more general, since they consider the case in which the extension of Δ may be other than the set of designated values.

[66] Peña [1984] defines an operator, F, on real numbers such that the value $F\alpha$ is 0 if that of α is greater than 0, and 1 otherwise; and then defines a conditional operator, $\alpha C \beta$, as $F\alpha \vee \beta$. The result is similar.

and so neither $(\alpha \wedge \neg\alpha) \to \beta$ nor $\alpha \to (\alpha \vee \neg\alpha)$ is valid, the system is not a relevant one since, e.g., $\models p \to (q \to q)$.

These facts may both be changed by modifying the semantics, by adding a class of non-normal worlds. Thus, an interpretation is a structure $\langle W, N, *, \nu\rangle$. The worlds in N are called *normal*; the worlds in $W - N$ (NN) are called *non-normal*. Truth conditions are the same as before, except that at non-normal worlds, the truth value of a conditional is arbitrary. Technically, ν assigns to every pair of world and propositional parameter a truth value, as before, but for every $w \in NN$ and every conditional $\alpha \to \beta$, it now also assigns $\alpha \to \beta$ a value at w. This provides the value of $\alpha \to \beta$ at non-normal worlds (non-recursively). Validity is defined as truth preservation at all *normal* worlds of all interpretations.

If one thinks of the conditionals as entailments, then the non-normal worlds are those where the facts of logic may be different. Thus, one may think of non-normal worlds as logically impossible situations.[67]

The system described is called H in Routley and Loparić [1978].[68] It is sound and weakly complete (i.e., theorem-complete) with respect to the following axiom system.

$\vdash \alpha \to \alpha$

$\vdash (\alpha \wedge \beta) \to \alpha \quad \vdash (\alpha \wedge \beta) \to \beta$

$\vdash \beta \to (\alpha \vee \beta) \quad \vdash \beta \to (\alpha \vee \beta)$

$\vdash \alpha \leftrightarrow \neg\neg\alpha$

$\vdash (\neg\alpha \vee \neg\beta) \leftrightarrow \neg(\alpha \wedge \beta)$

$\vdash (\neg\alpha \wedge \neg\beta) \leftrightarrow \neg(\alpha \vee \beta)$

$\vdash (\alpha \wedge (\beta \vee \gamma)) \to ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma))$

If $\vdash \alpha$ and $\vdash \alpha \to \beta$ then $\vdash \beta$

If $\vdash \alpha$ and $\vdash \beta$ then $\vdash \alpha \wedge \beta$

If $\vdash \alpha \to \beta$ and $\vdash \beta \to \gamma$ then $\vdash \alpha \to \gamma$

If $\vdash \alpha \to \beta$ then $\vdash \neg\beta \to \neg\alpha$

If $\vdash \alpha \to \beta$ and $\vdash \alpha \to \gamma$ then $\vdash \alpha \to (\beta \wedge \gamma)$

If $\vdash \alpha \to \gamma$ and $\vdash \beta \to \gamma$ then $\vdash (\alpha \vee \beta) \to \gamma$

Strong (i.e., deducibility-) completeness requires also the rules in disjunctive form.[69] The disjunctive form of the first is: $\vdash \alpha \vee \gamma$ and $\vdash (\alpha \to \beta) \vee \gamma$ then $\vdash \beta \vee \gamma$. The others are similar.[70]

[67]For a further discussion of non-normality, see Priest [1992].

[68]There are several other systems in the vicinity here. Some are obtained by varying the conditions on $*$. Others, sometimes called the Arruda - da Costa P systems, are obtained by retaining the positive logic and adding a non-truth-functional negation. For details, see Routely and Loparić [1978].

[69]Which are known to be admissible anyway.

[70]A sound and complete natural deduction system is an open question.

Soundness is proved as usual. The (strong) completeness proof is as follows. We first show by induction on proofs that if $\alpha \vdash \beta$ then $\alpha\vee\gamma \vdash \beta\vee\gamma$. It quickly follows that if $\alpha \vdash \gamma$ and $\beta \vdash \gamma$ then $\alpha \vee \beta \vdash \gamma$. Now suppose that $\Gamma \nvdash \alpha$. Extend Γ to a prime theory, Θ, with the same property, as in 4.3. Call a set Δ a Θ-*theory* if it is prime, closed under adjunction, and $\beta \rightarrow \gamma \in \Theta \Rightarrow (\beta \in \Delta \Rightarrow \gamma \in \Delta)$. Note that Θ is a Θ-theory. Define the interpretation $\langle W, N, *, \nu\rangle$, where W is the set of Θ-theories; $N = \{\Theta\}$, $\beta \in \Delta^*$ iff $\neg\beta \notin \Delta$ (which is well-defined). If $\Delta \in NN$ then $\nu_\Delta(\beta \rightarrow \gamma) = 1$ iff $\beta \rightarrow \gamma \in \Delta$; and for all Δ:

$$\nu_\Delta(p) = 1 \text{ iff } p \in \Delta$$

Once it can be shown that this condition carries over to all formulas, the result follows as usual. This is proved by induction. The only difficult case concerns $\rightarrow$ when $\Delta = \Theta$. From right to left, the result follows from the definition of W. From left to right, the result follows from the following lemma. If $\beta \rightarrow \gamma \notin \Theta$ then there is a Θ-theory, Δ, such that $\beta \in \Delta$ and $\gamma \notin \Delta$. To prove this, we proceed essentially as in 4.3, except that $\Sigma \vdash \Pi$ is redefined. Let $\Theta_\rightarrow$ be the set of conditionals in Θ; then $\Sigma \vdash \Pi$ is now taken to mean that there are $\sigma_1, ..., \sigma_n \in \Sigma$ and $\pi_1, ..., \pi_m \in \Pi$ such that $\Theta_\rightarrow \vdash (\sigma_1 \wedge ... \wedge \sigma_n) \rightarrow (\pi_1 \vee ... \vee \pi_m)$. Now set $\Sigma = \{\beta\}$, and $\Pi = \{\beta\}$, and proceed as in 4.3. The rest of the details are left as a (lengthy) exercise.[71]

If we add the Law of Excluded Middle to the axiom system:

$$\vdash \alpha \vee \neg\alpha$$

we obtain a logic that we will call HX. In virtue of the discussion in 4.7, one might suppose that this would be sound and complete if we add the condition: for all w, and parameters, p, $1 = \nu_w(p)$ or $0 = \nu_{w^*}(p)$. This condition indeed makes $\alpha \vee \neg\alpha$ true in all worlds; but for just that reason, it also verifies the irrelevant $\beta \rightarrow (\alpha \vee \neg\alpha)$. To obtain HX, we place this constraint on just normal worlds. The semantics are then just right, as may be checked. For further details, see Routley and Loparić [1978]. Since normal worlds are now, in effect, LP interpretations, HX verifies all the logical truths of LP and so of classical logic.

A feature of this system is that substitutivity of equivalents breaks down. For example, as is easy to check, $p \leftrightarrow q \not\models (r \rightarrow p) \leftrightarrow (r \rightarrow q)$. This can be changed by taking the valuation function to work on propositions (i.e., set of worlds), rather than formulas.[72] The most significant feature of semantics of this kind is that there are no principles of inference that employ nested

[71] Details can be found in Priest and Sylvan [1992].
[72] For details see Priest [1992].

conditionals in an essential way. This is due entirely to the anarchic nature of non-normal worlds. In effect, *any* breakdown of logic is countenanced.

One way of putting a little order into the anarchy without destroying relevance, proposed by Routley and Meyer,[73] is by employing a ternary relation, R, to give the truth conditions of conditionals at non-normal worlds. An interpretation is now of the form $\langle W, N, R, *, \nu\rangle$. All is as before, except that ν no longer gives the truth values of conditionals at non-normal worlds. Rather, for any $w \in NN$, the truth conditions are:

$$\nu_w(\alpha \to \beta) = 1 \text{ iff for all } x, y \in W, Rwxy \Rightarrow (\nu_x(\alpha) = 1 \Rightarrow \nu_y(\beta) = 1)$$

Note that this is just the standard condition for strict implication, except that the worlds of the antecedent (x) and the consequent (y) have become distinguished. What, exactly, the ternary relation, R, means, is still a matter for philosophical deliberation. Validity is again defined as truth preservation at all normal worlds.

These semantics give the basic system of affixing relevant logic, B. An axiom system therefor can be obtained by replacing the last two rules for H by the corresponding axioms:

$$\vdash ((\alpha \to \beta) \wedge (\alpha \to \gamma)) \to (\alpha \to (\beta \wedge \gamma))$$
$$\vdash ((\alpha \to \gamma) \wedge (\beta \to \gamma)) \to ((\alpha \vee \beta) \to \gamma)$$

and adding a rule that ensures replacement of equivalents:

$$\text{If} \vdash \alpha \to \beta \text{ and} \vdash \gamma \to \delta \text{ then} \vdash (\beta \to \gamma) \to (\alpha \to \delta)$$

The soundness and completeness proofs generalise those for H. Details can be found in Priest and Sylvan [1992].

We may form the system BX proof theoretically by adding the Law of Excluded Middle. Semantically, we proceed as with H, placing the appropriate condition on normal worlds.

As with modal logics, stronger logics can be obtained by placing conditions on the accessibility relation, R. In this way, most of the logics in the Anderson/Belnap family can be generated. Details can be found in Restall [1993]. The strongest of these is the logic R, an axiom system for which is as follows:

$$\vdash \alpha \to \alpha$$
$$\vdash (\alpha \to \beta) \to ((\beta \to \gamma) \to (\alpha \to \gamma))$$
$$\vdash \alpha \to ((\alpha \to \beta) \to \beta)$$

[73] Initially, this was in Routley and Meyer [1973]. For further discussion of all the following, see the article on Relevent Logic in this volume of the *Handbook*.

$\vdash (\alpha \to (\alpha \to \beta)) \to (\alpha \to \beta)$
$\vdash (\alpha \wedge \beta) \to \beta, \quad \vdash (\alpha \wedge \beta) \to \alpha$
$\vdash ((\alpha \to \beta) \wedge (\alpha \to \gamma)) \to (\alpha \to (\beta \wedge \gamma))$
$\vdash \beta \to (\alpha \vee \beta), \quad \vdash \alpha \to (\alpha \vee \beta)$
$\vdash ((\alpha \to \gamma) \wedge (\beta \to \gamma)) \to ((\alpha \vee \beta) \to \gamma)$
$\vdash (\alpha \wedge (\beta \vee \gamma)) \to ((\alpha \wedge \beta) \vee \gamma)$
$\vdash (\alpha \to \neg\beta) \to (\beta \to \neg\alpha)$
$\vdash \neg\neg\alpha \to \alpha$

with the rules of adjunction and *modus ponens.*

The equivalence between the Dunn 4-valued semantics and the Routley $*$ operation that we noted in 4.7 suggests another way of obtaining an intensional conditional connective. In the simplest case, an interpretation is a structure $\langle W, v\rangle$ where W is a set of worlds and ν is an evaluation of the parameters at worlds, but this time it is a Dunn 4-valued interpretation. The truth conditions for the basic language are as in 4.6, except that they are relativised to worlds. Thus, using the functional notation:

$1 \in \nu_w(\neg\alpha)$ iff $0 \in \nu_w(\alpha)$
$0 \in \nu_w(\neg\alpha)$ iff $1 \in \nu_w(\alpha)$

$1 \in \nu_w(\alpha \wedge \beta)$ iff $1 \in \nu_w(\alpha)$ and $1 \in \nu_w(\beta)$
$0 \in \nu_w(\alpha \wedge \beta)$ iff $0 \in \nu_w(\alpha)$ or $0 \in \nu_w(\beta)$

$1 \in \nu_w(\alpha \vee \beta)$ iff $1 \in \nu_w(\alpha)$ or $1 \in \nu_w(\beta)$
$0 \in \nu_w(\alpha \vee \beta)$ iff $0 \in \nu_w(\alpha)$ and $0 \in \nu_w(\beta)$

The natural truth and falsity conditions for $\to$ are:

$1 \in \nu_w(\alpha \to \beta)$ iff for all $w' \in W, (1 \in \nu_{w'}(\alpha) \Rightarrow 1 \in \nu_{w'}(\beta))$
$0 \in \nu_w(\alpha \to \beta)$ iff for some $w' \in W$, $1 \in \nu_{w'}(\alpha)$ and $0 \in \nu_{w'}(\beta)$

These semantics do not validate the undesirable:

$$\frac{\alpha \to \beta \qquad \neg(\alpha \to \beta)}{\gamma}$$

as their $*$ counterparts do. But they are still not relevant. Relevant logics can be obtained by adding a class of non-normal worlds. The semantic values of conditionals at these may either be arbitrary, as with H, or, as with B, we may employ a ternary relation and give the conditions as follows:

$1 \in \nu_w(\alpha \to \beta)$ iff for all $x, y \in W$, $Rwxy \Rightarrow (1 \in \nu_x(\alpha) \Rightarrow 1 \in \nu_y(\beta))$
$0 \in \nu_w(\alpha \to \beta)$ iff for some $x, y \in W$, $Rwxy, 1 \in \nu_x(\alpha)$ and $0 \in \nu_y(\beta)$

As usual, extra conditions may be imposed on R. This construction produces a family of relevant logics distinct from the usual ones, and one that has not been studied in great detail. One way in which it differs from the more usual ones is that contraposition of the conditional fails, though this can be rectified by modifying the truth conditions for $\to$ by adding the clause: 'and $0 \in \nu_{w'}(\beta) \Rightarrow 0 \in \nu_{w'}(\alpha)$' (or in the case of non-normal worlds employing a ternary relation: 'and $0 \in \nu_x(\beta) \Rightarrow 0 \in \nu_y(\alpha)$'). A more substantial difference concerns negated conditionals. Because of the falsity conditions of the conditional, all logics of this family validate $\alpha \wedge \neg\beta \models \neg(\alpha \to \beta)$. This is a natural enough principle, but absent from many of the logics obtained using the Routley $*$.

The more usual relevant logics can be obtained with the 4-valued semantics, but only by using some *ad hoc* device or other, such as an extra accessibility relation, or allowing only certain classes of worlds. For details, see Routley [1984] and Restall [1995].

5.6 $\to$ as $\leq$

There is a very natural way of employing any algebra which has an ordering relation to give a semantics for conditionals. One may think of the members of the algebra as propositions, or as Fregean senses. The relation $\leq$ on the algebra can be thought of as an entailment relation, and it is then natural to take $\alpha \to \beta$ to hold in some interpretation, ν, iff $\nu(\alpha) \leq \nu(\beta)$. The problem, then, is to express the thought that $\alpha \to \beta$ holds in algebraic terms. We obviously need an algebraic operator, $\to$, corresponding to the connective; but how is one to express the idea that $a \to b$ holds when the algebra may have no maximal element?

A way to solve this problem for De Morgan algebras is to employ a designated member of the lattice, e, and take the things that hold in the algebra to be those whose values are $\geq e$.[74] While we are introducing new machinery, it is also useful algebraically to introduce another binary (groupoid) operator, $\circ$, often called 'fusion', whose significance we will come back to in a moment. We may also enrich the basic language to one containing a constant, e, and an operator, $\circ$, expressing the new algebraic features.

Thus, following Meyer and Routley [1972], let us call the structure $\mathcal{A} = \langle \mathcal{D}, e, \to, \circ \rangle$ a *De Morgan groupoid* iff $\mathcal{D}$ is a De Morgan algebra, $\langle A, \wedge, \vee, \neg \rangle$, and for any $a, b, c \in A$:

[74] A different way is to let T be a prime filter on the lattice, thought of as the set of all true propositions. We can then require that $a \to b \in T$ iff $a \leq b$. For details, see Priest [1980].

$e \circ a = a$
$a \circ b \leq c$ iff $a \leq b \rightarrow c$
if $a \leq b$ then $a \circ c \leq b \circ c$ and $c \circ a \leq c \circ b$
$a \circ (b \vee c) = (a \circ b) \vee (a \circ c)$ and $(b \vee c) \circ a = (b \circ a) \vee (c \circ a)$

The first of these conditions ensures that e is a left identity on the groupoid. (Note that the groupoid may not be commutative.) And it, together with the second, ensure that $a \leq b$ iff $e \leq a \rightarrow b$. The third and fourth ensure that $\circ$ respects the lattice operations in a certain sense. The sense is question in that of a sort of conjunction, and this makes it possible to think of fusion as a kind of intensional conjunction.

An inference, $\alpha_1, ..., \alpha_n/\beta$, is algebraically valid iff for every homomorphism, ν, into a De Morgan groupoid, $\nu(\alpha_1 \wedge ... \wedge \alpha_n) \leq \nu(\beta)$, i.e., $e \leq \nu((\alpha_1 \wedge ... \wedge \alpha_n) \rightarrow \beta))$.[75]

These semantics are sound and complete with respect to the relevant logic B of 5.5. Soundness is shown in the usual way, and completeness can be proved, as in 4.8, by constructing the Lindenbaum algebra, and showing that it is a De Morgan groupoid.

Stronger logics can be obtained, as usual, by adding further constraints. The condition: $e \leq a \vee \neg a$ gives the law of excluded middle (and all classical tautologies). Additional constraints on $\circ$ give the stronger logics in the usual relevant family, including R. Details of all the above can be found in Meyer and Routley [1972] (who also show how to translate between algebraic and world semantics).[76]

Before leaving the topic of conditionals in algebraic paraconsistent logics, a final comment on dual intuitionist logic. Goodman [1981] proves that in this logic there is no conditional operator (i.e., operator satisfying *modus ponens*) that can be defined in terms of $\vee, \wedge$ and $\neg$; and draws somewhat pessimistic conclusions from this concerning the usefulness of the logic. Such pessimism is not warranted, however. Exactly the same is true in relevant logic; this does not mean that a conditional operator cannot be added to the basic language. And as Mortensen notes,[77] given any algebraic structure with top ($\top$) and bottom ($\bot$) elements, the following conditions can always be used to define a conditional operator:

$$\begin{aligned} \nu(\alpha \rightarrow \beta) &= \top \quad \text{if } \nu(\alpha) \leq \nu(\beta) \\ &= \bot \quad \text{otherwise} \end{aligned}$$

[75] A different notion of validity can be formulated using fusion thus: $\nu(\alpha_1 \circ ... \circ \alpha_n) \leq \nu(\beta)$, i.e., $e \leq \nu((\alpha_1 \rightarrow (... \rightarrow (\alpha_n \rightarrow \beta)...)$.

[76] See also Brink [1988]. A rather different algebraic approach which produces a relevant logic is given in Avro'n [1990]. This maintains an ordered structure, but dispenses with the lattice. The result is a logic closely related to the intensional fragment of RM.

[77] Mortensen [1995], p. 95.

Though this particular conditional is not suitable for robust paraconsistent purposes since it satisfies: $\alpha \rightarrow \beta, \neg(\alpha \rightarrow \beta) \models \gamma$.

5.7 *Decidability*

Before we leave the topic of propositional logics, let me review, briefly, the question of decidability for the logics that we have looked at. Unsurprisingly, most (though not all) are decidable, as the following decision procedures indicate. As will be clear, in many cases the procedures actually given could be greatly optimised.

Any filter logic is decidable if the filter is. Given any inference, we can effectively find the set of all inferences of which it is a uniform substitution instance. Provided that the filter is decidable, we can test each of these for prevalidity. If any of them is valid, the original inference is valid; otherwise not.

Smiley's filter is clearly decidable. So is Tennant's semantic filter. Given an inference with finite sets of premises and conclusions, Σ and Π, respectively, we can test the inference for classical validity. We may then test the inferences for all subsets of Σ and Π. (There is only a finite number of these.) If the original inference is valid, but its subinferences are not, it passes the test; otherwise not. Tennant's proof theory of 5.1 is also decidable. Anything provable has a Cut-free proof (since Cut is not a rule of proof). Decidability then follows as it does in the case of classical logic.

Turning to non-adjunctive logics: Jaśkowski's discursive logic is decidable; we may simply translate an inference into the corresponding one concerning $S5$, and use the $S5$ decision procedure for this. The same obviously goes for any generalisation, provided only that the underlying modal logic is decidable.

Rescher and Manor's logic is decidable in the obvious way. Given any finite set of premises, we can compute all its subsets, the classical consistency of each of these, and hence determine which of the sets are maximally consistent. Once we have these, we can determine if any of them classically entails the conclusion. Similar comments apply to Schotch and Jennings' logic. Given any premise set, we can compute all its partitions, and so determine its level. For every partition of that size, we can test to see if one of its members classically entails the conclusion.

Non-truth-functional logics are also decidable by a simple procedure. Given an inference, we consider the set of all subformulas of the sentences involved (which is finite). We then consider all mappings from these to $\{0, 1\}$, the set of which is also finite. For each of these we go through and test whether it satisfies the appropriate constraints in the obvious way. Throwing away all those that do not, we see whether the conclusion holds in all that remain.[78]

All finite many-valued logics are decidable by truth-tables. The infinite valued Łukasiewicz logics (and so their Kotas and da Costa augmentations) are not, in general, even axiomatisable, let alone decidable. (See Chang [1963].) This leaves RM. If there is a counter-model for an RM inference, there must be a number of maximum absolute value employed. Ignoring all the numbers in the model whose absolute size is greater than this gives a finite counter-model. Hence, RM has the finite model property. As is well known, any axiomatisable theory with this property is decidable. (Enumerate the theorems and the finite models simultaneously. Eventually we must find either a proof of a countermodel.)

Dual intuitionist logic is decidable since intuitionist logic is. We just compute the dual inference and test it with the intuitionist procedure.

This just leaves the logics of the relevant family. As we saw, the semantics of these can take either a world form or an algebraic form. The question of decidability here is the hardest and most sensitive. The weaker logics in the family are decidable, and can be shown to be so by semantic methods (such as filtration arguments) and/or proof theoretic ones (such as Gentzenisation plus Cut elimination).[79] The stronger ones, such as R, are not. Urquhart's [1984] proof of this fact contains one of the few applications of geometry to logic. A crucial principle in this context would seem to be contraction: $(\alpha \to (\alpha \to \beta)) \to (\alpha \to \beta)$ (or various equivalent forms, such as $(\alpha \wedge (\alpha \to \beta)) \to \beta)$. Speaking very generally, systems without this principle are decidable; systems with it are not.

6 QUANTIFIERS

The novelty of paraconsistent logic lies, it is fair to say, almost entirely at the propositional level. However, if a logic is to be applied in any serious way, it must be quantificational. Most of the paraconsistent logics that we have considered extend in straightforward ways to quantified logics. In this section I will indicate how. Let us suppose that the propositional language is now augmented to a language, L, with predicates, constants, variables and the quantifiers $\forall$ and $\exists$ in the usual way. I will let the adicity of a predicate be shown by the context. Propositional parameters can be identified with predicates of adicity 0. I will write $\alpha(x/t)$ to mean the result of substituting the term t for all free occurrences of x, any bound variables in α having been relabelled, if necessary, to avoid clashes.

I will reserve the word 'sentence' for formulas without free variables. I will always define validity for inferences containing only sentences, though the accounts could always be extended to ones employing all formulas, in

[78] For the method applied to the da Costa systems, see da Costa and Alves [1977].

[79] See, respectively, Routley *et al.* [1982], sect. 5.9, and Brady [1991].

standard ways. Where quantifiers have an objectual interpretation, and the set of objects is D, I will assume—for the rest of this essay—that the language has been augmented by a set of constants in such a way that each member of the domain has a name. In particular, I will always assume that the names are the members of D themselves, and that each object names itself. This assumption is never essential, but it simplifies the notation.

6.1 Filter and Non-adjunctive Logics

In filter logics, we may simply take the filter to be a relation on the extended language. Smiley's filter works equally well, for example, when the notion of classical logical truth employed is that for first order, not propositional, logic. Similarly for Tennant's. In his case (without the conditional operator), the semantics are sound and complete with respect to the sequent calculus of 4.1 for the basic language, together with the usual rules for the quantifiers:

$$\frac{\Gamma : \alpha(x/c), \Delta}{\Gamma : \forall x\alpha, \Delta} \qquad \frac{\Gamma, \alpha : \Delta}{\Gamma, \forall x\alpha : \Delta}$$

$$\frac{\Gamma : \alpha, \Delta}{\Gamma : \exists x\alpha, \Delta} \qquad \frac{\Gamma, \alpha(x/c) : \Delta}{\Gamma, \exists x\alpha : \Delta}$$

where in the first and last of these, c does not occur in any formula in Γ or Δ. For proofs, see Tennant [1984]. (With the conditional operator added, the situation is different, as we saw in 5.1.)

Non-adjunctive logic accommodates quantifiers in an obvious way. Consider discursive logic. An inference in the quantified language is discursively valid iff $\Diamond\Sigma \models_{CS5} \Diamond\alpha$, where $CS5$ is constant-domain quantified $S5$. Clearly, any other quantified modal logic could be used to generalise this notion.[80]

Rescher and Manor's approach and Schotch and Jennings' also generalise in the obvious way, the classical notion of propositional consequence involved being replaced by the classical first-order notion. In the quantificational case, the usefulness of these logics is moot, since the computation of classically maximally consistent sets of premises, or partitions, is highly non-effective.

In all these logics, except Smiley's, the set of logical truths (in the appropriate vocabulary) coincides with that of classical quantifier logic; hence these logics are undecidable.[81]

---

[80] For details of quantified modal logic, see the article on that topic in this *Handbook*.

[81] I do not know whether Smiley's logic is decidable, though I assume that it is not.

6.2 Positive-plus Logics

Let us turn now to the logics that augment classical or intuitionist positive logic with a non-truth-functional negation. Since the semantics of these are not truth functional, the most natural quantifier semantics are not objectual, but substitutional. Let me illustrate this with the simplest non-truth-functional logic, with a classical conditional operator, but no semantic constraints on negation. Extensions of this to other cases are left as an exercise.

An interpretation is a pair $\langle C, \nu \rangle$. C is a set of constants, and L_C is the language L augmented by the constants C. ν is a map from the sentences of L_C to $\{1, 0\}$ satisfying the same conditions as in the propositional case, together with:

$\nu(\forall x\alpha) = 1$ iff for every constant of L_C, c, $\nu(\alpha(x/c)) = 1$
$\nu(\exists x\alpha) = 1$ iff for some constant of L_C, c, $\nu(\alpha(x/c)) = 1$

An inference is valid iff it is truth-preserving in all interpretations.

The semantics are sound and complete with respect to the quantifier rules:

$$\forall I \quad \frac{\beta \vee \alpha(x/c)}{\beta \vee \forall x\alpha}$$

provided that c does not occur in β, or in any undischarged assumption on which the premise depends.

$$\forall E \quad \frac{\forall x\alpha}{\alpha(x/c)}$$

$$\exists I \quad \frac{\alpha(x/c)}{\exists x\alpha}$$

$$\exists E \quad \frac{\exists x\alpha \qquad \begin{matrix} \overline{\alpha(x/c)} \\ \vdots \\ \beta \end{matrix}}{\beta}$$

provided that c does not occur in β or in any undischarged assumption in the subproof.

Soundness is proved by a standard recursive argument. For completeness, call a theory, Σ, *saturated* in a set of constants, C, iff:

$\exists x\alpha \in \Sigma$ iff for some $c \in C$, $\alpha(x/c) \in \Sigma$

$\forall x\alpha \in \Sigma$ iff for every $c \in C$, $\alpha(x/c) \in \Sigma$

It is easy to check that if Δ is a prime theory, saturated in C, then $\langle C, \nu\rangle$ is an interpretation, where ν is defined by: $\nu(\alpha) = 1$ iff $\alpha \in \Delta$.

It remains to show that if $\Sigma \not\vdash \alpha$ then Σ can be extended to a prime theory, Δ, saturated in some set of constants, C, with the same property; and the result follows as in the propositional case, using Δ to define the interpretation.

To show this, we augment the language with an infinite set of new constants, C, and then extend the proof of 4.3 as follows. Enumerate the formulas of L_C: β_0, β_1,... If $\forall x\beta$ or $\exists x\beta$ occurs in the enumeration, and the constant c does not occur in any preceding formula, we will call $\beta(x/c)$ a *witness.* Now, we run through the enumeration, as before, but this time, if we throw $\exists x\beta$ into the Σ side, we also throw in a witness; and if we throw $\forall x\beta$ into the Π side, we also throw in a witness. In proving that $\Sigma_n \not\vdash \Pi_n$, the only novelty is when a witness is present; and these can be ignored, by $\exists E$ on the left, and $\forall I$ on the right. The rest of the proof is as in 4.3. The saturation of Δ in C follows from deductive closure and construction.

I observe that all the logics in this family contain positive classical quantifier logic, and so are undecidable.

6.3 Many-valued Logics

Most of the many-valued logics with numerical values that we considered in 4.5 and 5.4 had two particular properties. First, the truth value of a conjunction [disjunction] is the minimum [maximum] of the values of the conjuncts [disjuncts]. Second, the set of truth values is closed under greatest lower bounds (glbs) and least upper bounds (lubs), i.e., if $Y \subseteq X$ then $\text{glb}(Y) \in X$ and $\text{lub}(Y) \in X$. Any such logic can be extended to a quantified logic in a very natural way, merely by treating $\forall$ and $\exists$ as the "infinitary" generalisations of conjunction and disjunction, respectively.

Specifically, a quantifier interpretation adds to the propositional machinery, the pair $\langle D, d\rangle$ where D is a non-empty domain of objects, d maps every constant into D, and if P is an n-place predicate, d maps P to a function from n-tuples of the domain into the set of truth-values. Every sentence, α, can now be assigned a truth value, $\nu(\alpha)$, in the natural way. For atomic sentences, $Pc_1...c_n$:

$$\nu(Pc_1...c_n) = d(P)\,\langle d(c_1)...d(c_n)\rangle$$

The truth conditions for propositional connectives are as in the propositional logic. The truth conditions for the quantifiers are:

$$\nu(\forall x\alpha) = \text{glb}\{\alpha(x/c); c \in D\}$$
$$\nu(\exists x\alpha) = \text{lub}\{\alpha(x/c); c \in D\}$$

Validity is defined in terms of preservation of designated values, as in the propositional case.

I will make just a few comments about what happens when these definitions are applied to the many-valued logics we have looked at. The quantified finite-valued logics of 4.5 all collapse into quantified *LP* (which we will come to in the next section), as extensions of the arguments given there, show. For a general theory of quantified finitely-many-valued logics, see Rosser and Turquette [1952]. Quantified *RM* we will come to in a later section. Infinite-valued Łukasiewicz logics are proof-theoretically problematic. For a start, standard quantifier rules may break down. In particular, $\forall x\alpha$ may be undesignated, even though each substitution instance is designated. Thus, $\forall I$ may fail. (Similarly for existential quantification.) Worse, as for their propositional counterparts, such logics are not even axiomatisable in general.[82]

6.4 LP and FDE

The technique of extending a many-valued logic to a quantified one can be put in a slightly different, and possibly more illuminating, way for the logics with relational semantics, *LP* and *FDE*. An interpretation, $\mathcal{I}$, is a *pair*, $\langle D, d\rangle$, where D is the usual domain of quantification, d is a function that maps every constant into the domain, and every n-place predicate into a pair, $\langle E_P, A_P\rangle$, each member of which is a subset of the set of n-tuples of D, D^n. E_P is the *extension* of P; A_P is the *anti-extension*. For *LP* interpretations, we require, in addition, that $E_P \cup A_P = D^n$. Truth values are now assigned to sentences in accord with the following conditions. For atomic sentences:

$$1 \in \nu(Pc_1...c_n) \text{ iff } \langle d(c_1), ..., d(c_n)\rangle \in E_P$$
$$0 \in \nu(Pc_1...c_n) \text{ iff } \langle d(c_1), ..., d(c_n)\rangle \in A_P$$

Truth/falsity conditions for connectives are as in the propositional case; and for the quantifiers:

$$1 \in \nu(\forall x\alpha) \text{ iff for every } c \in D,\ 1 \in \nu(\alpha(x/c))$$
$$0 \in \nu(\forall x\alpha) \text{ iff for some } c \in D,\ 0 \in \nu(\alpha(x/c))$$

$$1 \in \nu(\exists x\alpha) \text{ iff for some } c \in D,\ 1 \in \nu(\alpha(x/c))$$

[82] See Chang [1963] for details.

$0 \in \nu(\exists x\alpha)$ iff for every $c \in D$, $0 \in \nu(\alpha(x/c))$

An inference is valid iff it is truth-preserving in all interpretations. It should be noted that if for every predicate, P, E_P and A_P are exclusive and exhaustive, we have an interpretation of classical first order logic. All classical interpretations are therefore FDE (and LP) interpretations.

These semantics are sound and complete if we add to the rules for LP or FDE, the rules $\forall I$, $\forall E$, $\exists I$ and $\exists E$, plus:

$$\frac{\forall x\neg\alpha}{\overline{\neg\exists x\alpha}} \qquad \frac{\exists x\neg\alpha}{\overline{\neg\forall x\alpha}}$$

Soundness is established by the usual argument. For completeness, suppose that $\Sigma \not\vdash \alpha$. Extend Σ to a set Δ, which is prime, deductively closed and saturated in a set of new constants, such that $\Delta \not\vdash \alpha$, as in 6.2. Then define an interpretation $\langle D, d\rangle$ where D is the set of constants of the extended language, d maps any constant to itself, and for any predicate, P, its extension and anti-extension are defined as follows:

$\langle c_1, ..., c_n\rangle \in E_P$ iff $Pc_1...c_n \in \Delta$
$\langle c_1, ..., c_n\rangle \in A_P$ iff $\neg Pc_1...c_n \in \Delta$

We now establish that for all formulas, α :

$1 \in \nu(\alpha)$ iff $\alpha \in \Delta$
$0 \in \nu(\alpha)$ iff $\neg\alpha \in \Delta$

The argument is a routine induction. Here are the cases for $\forall$.

$\forall x\alpha \in \Delta$	$\Leftrightarrow$	for all c, $\alpha(x/c) \in \Delta$	saturation
	$\Leftrightarrow$	for all c, $1 \in \nu(\alpha(x/c))$	induction hypothesis
	$\Leftrightarrow$	$1 \in \nu(\forall x\alpha)$	truth conditions of $\forall$

$\neg\forall x\alpha \in \Delta$	$\Leftrightarrow$	$\exists x\neg\alpha \in \Delta$	quantifier rules
	$\Leftrightarrow$	for some c, $\neg\alpha(x/c) \in \Delta$	saturation
	$\Leftrightarrow$	for some c, $0 \in \nu(\alpha(x/c))$	induction hypothesis
	$\Leftrightarrow$	$0 \in \nu(\forall x\alpha)$	truth conditions of $\forall$
	$\Leftrightarrow$	$1 \in \nu(\neg\forall x\alpha)$	truth conditions of $\neg$

The monotonicity property of the propositional logics LP and FDE carries over to the quantified case. If $\mathcal{I}_1$ and $\mathcal{I}_2$ are any interpretations, with truth value assignments ν_1 and ν_2, define $\mathcal{I}_1 \leq \mathcal{I}_2$ to mean that $\mathcal{I}_1$ and $\mathcal{I}_2$ have the same domain, and for every predicate, P, the extension (anti-extension) of P in $\mathcal{I}_1$ is a subset of the extension (anti-extension) of P in

$\mathcal{I}_2$. A simple induction shows that if $\mathcal{I}_1 \leq \mathcal{I}_2$ then for all formulas, α (in a language with a name for every member of the domain), $\nu_1(\alpha) \subseteq \nu_2(\alpha)$. As in 4.6, it follows that the set of logical truths of LP is exactly the same as that of classical first order logic. And FDE has no logical truths (just consider an interpretation that makes the extension and anti-extension of every predicate empty).

Since classical quantifier logic is not decidable, neither is quantified LP. If P is any n-place predicate, let P_{LEM} be the sentence: $\forall x_1 ... \forall x_n (Px_1 ... x_n \vee \neg Px_1 ... x_n)$. If P_{LEM} is true in an interpretation, then the extension and anti-extension of P, exhaust the n-tuples of the domain. If α is any formula, let α_{LEM} be the conjunction of all formulas of the form P_{LEM}, where P occurs in α. It follows that $\alpha_{LEM} \models \alpha$ in FDE iff $\models \alpha$ in LP. Hence, quantified FDE is undecidable too.

6.5 Relevant Logics

Turning to relevant logics, the issues are more complex. This is due to the fact that there are various approaches to these logics, the variety of the logics themselves, and the intrinsic complexities of the stronger logics.

Let us start with the world semantics. As we saw in 5.5, a world semantics for a relevant logic with the Routely operator is a structure $\langle W, N, *, \nu\ (, R)\rangle$, where W is a set of worlds, N is a subclass of normal worlds (the complement being NN), $*$ is the Routley operation (such that $w = w^{**}$), and ν assigns truth values to all propositional parameters at worlds. In the logic H, it also assigns values to conditionals at non-normal worlds. In stronger logics, the ternary relation R is present, and is used to specify the values of conditionals at non-normal worlds. When no constraints are placed on R, we have the logic B.

The simplest way of extending such semantics to those of a quantified language is by removing ν from the structure and adding a domain of quantification, D, and a denotation function d. d specifies a denotation for each constant (same at each world) and an extension for each n-place predicate at each world, $d_w(P) \subseteq D^n$. Truth conditions are given in the standard way. In particular, for the quantifiers:

$$\nu(\forall x\alpha) = 1 \text{ iff for every } c \in D,\ \nu(\alpha(x/c)) = 1$$
$$\nu(\exists x\alpha) = 1 \text{ iff for some } c \in D,\ \nu(\alpha(x/c)) = 1$$

An inference is valid iff it is truth-preserving in all *normal* worlds of all interpretations.[83]

[83]More complex semantics can be employed in the usual variety of ways employed in modal logic. (See the article on Quantified Modal Logic in this *Handbook*.) In particular, we might employ variable-domain semantics. This makes matters more complex.

Consider the following quantifier axioms and rules (where $\vdash$ is now taken to indicate universal closure):

$\vdash \forall x\alpha \rightarrow \alpha(x/c)$
$\vdash \alpha(x/c) \rightarrow \exists x\alpha$
$\vdash \alpha \wedge \exists x\beta \rightarrow \exists x(\alpha \wedge \beta)$ $\quad x$ not free in α
$\vdash \forall x(\alpha \vee \beta) \rightarrow (\alpha \vee \forall x\beta)$ $\quad x$ not free in β

If $\vdash \forall x(\alpha \rightarrow \beta)$ then $\vdash \exists x\alpha \rightarrow \beta$ $\quad x$ not free in β
If $\vdash \forall x(\alpha \rightarrow \beta)$ then $\vdash \alpha \rightarrow \forall x\beta$ $\quad x$ not free in α

It is easy to check that these axioms/rules are valid/truth-preserving for H. If they are added to the propositional axioms/rules for H, they are also complete. For the proof, see Routley and Loparić [1980].[84]

If we strengthen the two rules to conditionals (so that the first of these becomes $\vdash \forall x(\alpha \rightarrow \beta) \rightarrow (\exists x\alpha \rightarrow \beta)$, etc.) and add them to the rules for B, they are also sound and complete. The same is true for a number of the extensions of B, including BX. (For details, see Routley [1980a].) A notable exception to this fact is the system R. Though the system is sound, it is, perhaps surprisingly, not complete.[85] In fact, a proof-theoretic characterisation of constant domain quantified R is still an open problem. The axioms and rules are complete for the stronger semi-relevant system RM of 5.4.[86]

Since every relevant logic in the above family contains FDE, and this is undecidable, it follows that all the logics in this family are also undecidable.

6.6 Algebraic Logics

Given any algebraic logic, for which the appropriate algebraic structures are lattices, and in which conjunction and disjunction behave as lattice meet and join, there is, as with many-valued logics, a natural way to extend the machinery to quantifiers. An algebra is *complete* iff it is closed under least upper bounds ($\bigvee$) and greatest lower bounds ($\bigwedge$), i.e., if the domain of the algebra is A and $B \subseteq A$ then $\bigvee B \in A$ and $\bigwedge B \in A$. If $\mathcal{A}$ is any algebraic structure of the required kind, with domain A, then an interpretation is a triple $\langle \mathcal{A}, D, d\rangle$, where D is the domain of quantification, d maps every constant into D and every n-place predicate into a function from D^n into

(The philosophical gain, however, is dubious: world relativised quantifiers can always be defined in constant-domain semantics, provided we have an Existence predicate.)

[84]If one works with a free-variable notion of deducibility, as Routley and Loparić do, one also has to add the rule of universal generalisation: if $\vdash \alpha$ then $\vdash \forall x\alpha$.

[85]As Fine showed. Fine also produced a rather different semantics with respect to which it is complete. See Anderson *et al.* [1992], sects. 52 and 53.

[86]See Anderson *et al.* [1992], sect. 49.2.

A. Algebraic values are then assigned to all formulas in the usual way. In particular, for quantified sentences the conditions are:

$$\nu(\forall x\alpha) = \bigwedge\{\nu(\alpha(x/c)); c \in D\}$$
$$\nu(\exists x\alpha) = \bigvee\{\nu(\alpha(x/c)); c \in D\}$$

I will comment on this construction for only two kinds of algebras. The first is when $\mathcal{A}$ is a De Morgan groupoid, or strengthening thereof. In this case, the above semantics clearly give quantified relevant logics. Their relation to quantified relevant logics based on the intensional semantics has not, as far as I am aware, been investigated.

The second is where $\mathcal{A}$ is a dual intuitionist algebra. In this case, the semantics give a quantified logic that is dual to quantified intuitionist logic. For details, see Goodman [1981].[87]

6.7 *A Brief Look Back*

Now that we have surveyed a large number of paraconsistent logics up to a quantified level—some very briefly—it would seem appropriate to look back for a moment and put the systems into some sort of perspective.

The logics we have looked at fall roughly and inexactly into four categories: non-transitive logics, non-adjunctive logics, non-truth-functional logics and relevant logics. (The most interesting many-valued systems are zero degree relevant logic, FDE, or closely related to it, like LP, and so may be classed in this family.) The non-transitive logics seem to be good for extracting the essential juice out of classical inferences, but do not really take inconsistent semantic structures seriously. Non-adjunctive logics may be just what one needs for certain applications (e.g., inferences in a data base, where one would not necessarily want to infer $\alpha \wedge \neg\alpha$ from α and $\neg\alpha$); they also take inconsistent structure seriously, though conjoined contradictions are handled indiscriminately, which makes them unsuitable for many applications. Non-truth-functional logics contain the whole of classical (or at least intuitionist) positive logic, and so are useful when strong canons of positive reasoning are required. However, this very strength is a weakness when it comes to some important applications, as we shall see in connection with set theory. Undoubtedly the simplest and most robust paraconsistent logic is the logic LP. When conditional operators are required, the relevant logic BX is a good all-purpose paraconsistent logic. Its conditional operator is satisfactory for many purposes, but may be considered relatively weak. It may be strengthened to give stronger relevant logics; but this, too, may cause a problem for some applications, as we shall see.

[87]There is also a topos-theoretic account of quantification for dual intuitionistic logic. See Mortensen [1995], ch. 11.

7 OTHER EXTENSIONS OF THE BASIC APPARATUS

I now want to look at other extensions of the basic paraconsistent apparatus. One way or another, all the paraconsistent logics we have looked at can be extended appropriately. However, it is tedious to run through every case, especially when details are often obvious. Hence, I shall illustrate the extensions mainly with respect to just one logic. Since LP is simple and natural, it recommends itself for this purpose. I will comment on other logics occasionally, when there is a point to doing so.

7.1 Identity and Function Symbols

LP—and all the other logics with objectual semantics that we have looked at—can be extended to include function symbols and identity in the usual way. The denotation function, d, maps each n-place function symbol, f, to an n-place function on the domain. A denotation for every (closed) term, t, is then obtained by the usual recursive condition:

$$d(ft_1...t_n) = d(f)(d(t_1), ..., d(t_n))$$

With functional terms present, the quantifier rules of proof are extended to arbitrary (closed) terms in the usual way.

If we require the extension of the identity predicate to be $\{\langle x, x\rangle\,; x \in D\}$ then this is sufficient to validate the usual laws of identity:

$$\frac{}{t = t} \qquad \frac{t_1 = t_2 \quad \alpha(x/t_1)}{\alpha(x/t_2)}$$

This does not require identity statements to be consistent. In LP the anti-extension of identity is any set whose union with the extension exhausts D^2, and so a pair can be in both the extension and the anti-extension of the identity predicate. In other logics, negated identities can be taken care of by whatever mechanism is used for negation. The completeness proof for quantified LP can be extended to include function symbols and identity in the usual Henkin fashion.

I note that description operators can be added in the obvious ways, with the same panoply of options as in the classical case.[88]

7.2 Second-order Logic

Paraconsistent logics can also be extended to second order in the obvious ways. Consider LP. We add (monadic) second order variables, X, Y,... to

[88]See the article of Free Logics in this ***Handbook***.

the first-order language. Then, given an interpretation, $\langle D, d \rangle$, we extend the language to one, L_D, such that every member of the domain has a name, and for every pair E, A such that $E \cup A = D$ there is a predicate, P, with E and A as extension and anti-extension, respectively.[89] The truth/falsity conditions for the second order universal quantifier are then:

$1 \in \nu(\forall X \alpha)$ iff for every P in L_D, $1 \in \nu(\alpha(X/P))$
$0 \in \nu(\forall X \alpha)$ iff for some P in L_D, $0 \in \nu(\alpha(X/P))$

The truth/falsity conditions for the existential quantifier are the dual ones.

Appropriate monotonicity carries over to second order LP. Recall from 6.4 that if $\mathcal{I}_1$ and $\mathcal{I}_2$ are any interpretations, with truth value assignments ν_1 and ν_2, $\mathcal{I}_1 \leq \mathcal{I}_2$ means that $\mathcal{I}_1$ and $\mathcal{I}_2$ have the same domain, and for every predicate, P, the extension (anti-extension) of P in $\mathcal{I}_1$ is a subset of the extension (anti-extension) of P in $\mathcal{I}_2$. The same sort of induction as in the first-order case shows that if $\mathcal{I}_1 \leq \mathcal{I}_2$ then for all formulas, α, in L_D, $\nu_1(\alpha) \subseteq \nu_2(\alpha)$. (The predicates added in forming L_D have the same extension/anti-extension in both interpretations; and thus atomic sentences containing them satisfy the condition.)

In the second order case, and unlike the first order case, the logical truths of LP are distinct from their classical counterparts. For example, as is easy to check, in LP, $\models \exists X(Xa \wedge \neg Xa)$ (just consider the predicate which has D as both extension and anti-extension).[90] In fact, the logical truths of second order LP are inconsistent, since it is also a logical truth that $\forall X(Xa \vee \neg Xa)$, which is equivalent by quantifier rules and De Morgan to $\neg\exists X(Xa \wedge \neg Xa)$.

7.3 Modal Operators

All the logics may have modal operators added to them in one way or another. In the case of discursive logics, indeed, the semantics already provide for the possibility of alethic modal operators.

Adding modal operators to intensional logics where negation is handled by the Routley $*$ operator is very natural, but suffers problems similar to those we witnessed at the start of 5.5 in connection with the conditional. Suppose we take an intensional interpretation and give the modal operators the natural $S5$ conditions:

$\nu_w(\Box\alpha) = 1$ iff for every $w' \in W, \nu_{w'}(\alpha) = 1$

[89] This is the natural policy, since properties are characterised semantically by an extension/anti-extension pair. As in the classical case, there are other policies, e.g., where only predicates corresponding to some restricted class of properties are added.

[90] Second order FDE is constructed in the obvious way. The same sentence is a logical truth of this, showing that, unlike the first order case, it has logical truths.

$$\nu_w(\Diamond\alpha) = 1 \text{ iff for some } w' \in W, \nu_{w'}(\alpha) = 1$$

(or even N instead of W). Then the truth values of modalised statements are the same at all worlds. Hence, $\nu_w(\Box\alpha) = 1 \Leftrightarrow \nu_{w^*}(\Box\alpha) = 1 \Leftrightarrow \nu_w(\neg\Box\alpha) = 0$. Hence $\Box\alpha, \neg\Box\alpha \models \beta$, and so the logic is not suitable for serious paraconsistent purposes. The problem does not arise if we attempt a modal logic weaker than $S5$, for then the truth conditions of modal operators are given employing a binary accessibility relation in the usual way, and the truth values of modal statements will vary across worlds. But, at least for some purposes, an $S5$ modality is desirable.

These problems are avoided if we use the Dunn semantics for negation. The values of modalised formulas will still be the same at all worlds (in the $S5$ case), but we may now have both $\Box\alpha$ and $\neg\Box\alpha$ true at a world. I will illustrate, again, with respect to LP. Let us start with the case where the binary accessibility relation is arbitrary, the three-valued analogue of the modal system K.

An interpretation is now a structure $\langle W, R, \nu\rangle$, where W is a set of worlds; R is a binary relation on W; and for each parameter, p, $\nu_w(p) \in \{\{1\},\{1,0\},\{0\}\}$. Truth/falsity conditions for the propositional connectives are as in 5.5. The conditions for $\Box$ are:

$$1 \in \nu_w(\Box\alpha) \text{ iff for every } w' \text{ such that } wRw',\ 1 \in \nu_{w'}(\alpha)$$
$$0 \in \nu_w(\Box\alpha) \text{ iff for some } w' \text{ such that } wRw',\ 0 \in \nu_{w'}(\alpha)$$

and dually for $\Diamond$.[91]

It is easy to check that at every world of an interpretation $\Box\neg\alpha$ has the same truth value as $\neg\Diamond\alpha$, and dually. In fact, we can simply define $\Diamond\alpha$ as $\neg\Box\neg\alpha$, and will do this in what follows.

To obtain a proof-theoretic characterisation for the logic, we add to the rules for LP the following (chosen to make the completeness proof simple):

$$\frac{\begin{array}{c}\overline{\gamma}\\ \vdots\\ \beta\end{array} \qquad \Box\gamma}{\Box\beta} \qquad\qquad \frac{\begin{array}{c}\overline{\beta}\\ \vdots\\ \delta\end{array} \qquad \Diamond\beta}{\Diamond\delta}$$

where there are no other undischarged assumptions in the sub-proofs.

$$\frac{\Box(\beta \vee \delta)}{\Box\beta \vee \Diamond\delta} \qquad \frac{\Box\gamma \wedge \Diamond\beta}{\Diamond(\gamma \wedge \beta)}$$

[91] If a conditional operator is required, we may add a class of non-normal worlds—and maybe a ternary accessibility relation—and proceed as in 5.5.

$$\frac{\Box\gamma_1 \wedge ... \wedge \Box\gamma_n}{\Box(\gamma_1 \wedge ... \wedge \gamma_n)} \qquad \frac{\Diamond(\delta_1 \vee ... \vee \delta_n)}{\Diamond\delta_1 \vee ... \vee \Diamond\delta_n}$$

Soundness is easily checked. For completeness, suppose that $\Sigma \nvdash \alpha$. Extend Σ to a prime, deductively closed theory, Π, with the same property, as in 4.3. Define an interpretation, $\langle W, R, \nu\rangle$, where W is the set of prime deductively closed theories; $\Gamma R\Delta$ iff for all β:

$$\Box\beta \in \Gamma \Rightarrow \beta \in \Delta$$
$$\beta \in \Delta \Rightarrow \Diamond\beta \in \Gamma$$

and ν is defined by:

$$1 \in \nu_\Gamma(p) \text{ iff } p \in \Gamma$$
$$0 \in \nu_\Gamma(p) \text{ iff } \neg p \in \Gamma$$

All that remains is to show that these conditions extend to all formulas. Completeness then follows as usual. This is established by induction. The only difficult case is that for $\Box$, which requires the following two-part lemma.

If $\Box\beta \notin \Gamma$ then there is a $\Delta \in W$ such that $\Gamma R\Delta$ and $\beta \notin \Delta$. Proof: Let $\Gamma_\Box = \{\gamma; \Box\gamma \in \Gamma\}$ and $\Gamma_\Diamond = \{\delta; \Diamond\delta \notin \Gamma\}$. Then $\Gamma_\Box \nvdash \beta, \Gamma_\Diamond$, by the first and second pair of rules, and a bit of fiddling with the third. Extend $\Gamma_\Box$ to a prime, deductively closed set, Δ, with the same property, as in 4.3. The result follows.

If $\Diamond\beta \in \Gamma$ then there is a $\Delta \in W$ such that $\Gamma R\Delta$ and $\beta \in \Delta$. Proof: Let $\Gamma_\Box$ and $\Gamma_\Diamond$ be as before. Then $\Gamma_\Box, \beta \nvdash \Gamma_\Diamond$, by the first and second pair of rules, and a bit of fiddling with the third. Extend $\Gamma_\Box, \beta$ to a prime, deductively closed set, Δ, with the same property, as in 4.3. The result follows.

We can now prove the induction step for $\Box$:

$\Box\beta \in \Gamma$	$\Leftrightarrow$	$\forall\Delta$ s.t. $\Gamma R\Delta$, $\beta \in \Delta$	lemma in one direction definition of R in the other
	$\Leftrightarrow$	$\forall\Delta$ s.t. $\Gamma R\Delta$, $1 \in \nu_\Delta(\beta)$	induction hypothesis
	$\Leftrightarrow$	$1 \in \nu_\Gamma(\Box\beta)$	

$\neg\Box\beta \in \Gamma$	$\Leftrightarrow$	$\Diamond\neg\beta \in \Gamma$	definition of $\Diamond$
	$\Leftrightarrow$	$\exists\Delta(\Gamma R\Delta$ and $\neg\beta \in \Delta)$	lemma in one direction definition of R in the other
	$\Leftrightarrow$	$\exists\Delta(\Gamma R\Delta$ and $0 \in \nu_\Delta(\beta))$	induction hypothesis
	$\Leftrightarrow$	$0 \in \nu_\Gamma(\Box\beta)$	

Stronger modal logics can be obtained by placing conditions on R, and corresponding conditions on the proof theory. But even if we make R universal (so that for all x and y, xRy), and obtain the analogue of $S5$, we still do not get $\Box\alpha, \neg\Box\alpha \models \beta$. To see this, merely consider the interpretation with one world, w, which accesses itself; and where $\nu_w(p) = \{1, 0\}$ and $\nu_w(q) = \{0\}$. It is easy to check that $\nu_w(\Box p) = \nu_w(\neg\Box p) = \{1, 0\}$. Hence, $\Box p, \neg\Box p \not\models q$.

The same treatment can be given to temporal operators. If we take these, as usual, to be F and G for the future, and P and H for the past, then (three-valued) tense logic gives F and G the same truth conditions as $\Diamond$ and $\Box$, respectively; and P and H are the same, except that R is replaced by its converse, $\breve{R}$ (where $x\breve{R}y$ iff yRx). Appropriate soundness and completeness proofs for the case where R is arbitrary are obtained by modifying the alethic modal argument,[92] and stronger tense logics are obtained by adding conditions on R, in the usual way.[93]

Let me also mention conditional operators, $>$, of the Lewis/Stalnaker variety. These are modal (binary) operators, and can be given LP (or FDE) semantics in the same way that they are given a more usual semantics. For example, for the Stalnaker version, one extends interpretations with a selection function, $f(w, \alpha)$, thought of as selecting the nearest world to w where α is true. $\alpha > \beta$ is then true at w iff β is true at $f(w, \alpha)$. Details are left as a very non-trivial exercise.[94]

7.4 The Paraconsistent Importance of Modal Operators

Let me digress from the technical details to say a little about why modal operators are important in the context of paraconsistency. The reason is simply that so many of the natural areas where one might want to apply a paraconsistent logic involve them.

Take alethic modalities first. Even though one might not think that there are any true contradictions, one might still take them to be possible, in the sense of holding in some situations, such as fictional or counterfactual ones. Thus, one might hold that for some p, $\Diamond(p \wedge \neg p)$. This has a simple and obvious model in the above semantics. In this context, let me mention again the importance for counterfactual conditionals of worlds where the impossible holds; "impossible worlds" are just what one needs to evaluate such conditionals, according to the Lewis/Stalnaker semantics in whose direction I have just gestured.

Some have been tempted not just by the view that some contradictions

[92]See Priest [1982].

[93]See the article on Tense Logic in this ***Handbook***.

[94]For a paraconsistent theory of conditionals of this kind, and of many other modal operators, that employs the Routley $*$ to handle negation, see Routley [1989].

are possible, but by the view that *everything* is possible.[95] The valuation $\nu_{\{1,0\}}$ assigns every formula the value $\{1, 0\}$. (See 4.6). Hence, any interpretation that contains $\nu_{\{1,0\}}$ as one world will verify $\Diamond\alpha$, for all α, at any world that accesses it.[96] If we interpret the modal operators $\Box$ and $\Diamond$ as the deontic operators O (it is obligatory that) and P (it is permissible that), respectively, then the thesis that everything is possible becomes the nihilistic thesis that everything is permissible—what, according to Dostoevski, would be the case if there is no God.

Less exotically, standard deontic logic suffers badly from explosion.[97] Since in classical logic $\alpha, \neg\alpha \models \beta$ it follows that $O\alpha, O\neg\alpha \models O\beta$: if you have inconsistent obligations then you are obliged to do everything. This is surely absurd. People incur inconsistent obligations; this may give rise to legal or moral dilemmas, but hardly to legal or moral anarchy.[98] And one does not have to believe in dialetheism to accept this. Unsurprisingly, deontic explosion fails, given the semantics of the previous section: just consider the interpretation where there is a single world, w; R is universal; $\nu_w(p) = \{1, 0\}$ and $\nu_w(q) = \{0\}$. It is not difficult to check that $\nu_w(Op) = \nu_w(O\neg p) = \{1, 0\}$, whilst $\nu_w(Oq) = \{0\}$.

What is often taken to be the basic possible-worlds deontic logic (called KD by Chellas [1980], p. 131) makes matters even worse, by requiring that in an interpretation the accessibility relation be serial: for all x, there is a y such that xRy. This validates the inference $O\alpha/P\alpha$. It also validates the inference $O\neg\alpha/\neg O\alpha$. Hence we have, classically, $O\alpha, O\neg\alpha \models O\alpha \wedge \neg O\alpha \models \beta$; one who incurs inconsistent obligations renders the world trivial. Someone who believes that there are deontic dilemmas may just have to jettison the view that obligation entails permission, and so give up seriality. But on the above account one can retain seriality, and so both the above inferences; for $O\alpha \wedge \neg O\alpha \not\models \beta$, as the countermodel of the last paragraph shows.[99]

Another standard way of interpreting the modal operator $\Box$ is as an epistemic operator, K (it is known that), or a doxastic operator B (it is believed that). In these cases, classically, one would almost certainly want to put extra constraints on the accessibility relation, though what these should be might be contentious: all can accept reflexivity (xRx) for K (but not for B) since this validates $K\alpha \models \alpha$. Whether one would want transitivity ($(xRy \& yRz) \Rightarrow xRz$) is much more dubious for B and K, since this gives the

[95] E.g., Mortensen [1989].

[96] A similar, but slightly more complex, construction can be employed to the same effect if the logic has a conditional operator.

[97] See the article on Deontic Logic in this *Handbook* for details of Deontic Logic, including the possible-worlds approach.

[98] For further discussion, see Priest [1987], ch. 13.

[99] We have just been dealing with some of the "paradoxes of deontic logic". There are many of these. Arguably, all of them—or at least all the serious ones—are avoided by using a paraconsistent logic with a relevant conditional. See Routley and Routley [1989].

highly suspect $K\alpha \models KK\alpha$ and $B\alpha \models BB\alpha$. All this applies equally to the semantics of the previous section. Moreover, the paraconsistent semantics solve problems for doxastic logic of the same kind as for deontic logic. It is clear that people sometimes have inconsistent beliefs (if not knowledge). Standard semantics give $B\alpha, B\neg\alpha \models B\beta$. Yet patently someone may have inconsistent beliefs without believing everything.[100]

Observations such as this are particularly apt in the branch of *AI* known as knowledge representation, where it is common to use epistemic operators to model the information available to a computer. (See, e.g., a number of the essays in Halpern [1986].) Such information may well be inconsistent.

Finally, to tense operators. Whilst one does not have to be a dialetheist to hold that inconsistencies may be believed, obligatory, or true in some counterfactual situation, one does have to be, to believe that they were or will be true. Such views have certainly been held, however. Following Zeno, the whole dialectical tradition holds that contradictions arise in a state of change. To see one of the more plausible examples of this, just consider a state described by p which changes instantaneously at time t_0 to a state described by $\neg p$. What is the state of affairs at t_0? One answer is that at t_0, $p \wedge \neg p$ is true. Indeed, the contradictory state *is* the state of change.[101]

This can be modeled by the paraconsistent interpretation $\langle W, R, \nu \rangle$, where W is the set of real numbers (thought of as times); R is the standard ordering on the reals; and ν is defined by the condition:

$$\begin{aligned} \nu_t(p) &= \{1\} && \text{if } t < t_0 \\ &= \{1, 0\} && \text{if } t = t_0 \\ &= \{0\} && \text{if } t > t_0 \end{aligned}$$

It is easy to check that this interpretation verifies the inference: $p \wedge F\neg p / (p \wedge \neg p) \vee F(p \wedge \neg p)$, which we might call 'Zeno's Principle': change implies contradiction.

7.5 *Probability*

Probability is not a modal notion. But it, too, has paraconsistent significance. One of the most natural ways of constructing a paraconsistent probability theory is to extract one from a class of paraconsistent interpretations, in the manner of Carnap.[102]

[100] If you believe classical logic, then you might suppose that they are *rationally committed* to everything, but that is quite different. Even here, however, an explosive logic would seem to go astray. Dialetheism aside, situations such as the paradox of the preface, as well as more mundane things, would seem to show that one can be rationally committed to inconsistent propositions without being rationally committed to everything. See Priest [1987], sect. 7.4.

[101] See Priest [1982] and Priest [1987], ch. 11.

[102] See Carnap [1950].

A probabilistic interpretation is a pair, $\langle I, \mu \rangle$, where I is a class of interpretations for LP[103] and μ a finitely additive measure on I, that is, a function from subsets of I to non-negative real numbers such that:

$$\mu(\phi) = 0$$
$$\mu(X \cup Y) = \mu(X) + \mu(Y) \quad \text{if } X \cap Y = \phi$$

If α is any sentence, let $[\alpha] = \{\nu \in I; 1 \in \nu(\alpha)\}$. For reasons that we will come to, we also require that for all α, $\mu([\alpha]) \neq 0$. There certainly are such interpretations and measures. For example, let I be any finite class that contains the trivial interpretation, $\nu_{\{1,0\}}$, where all sentences are true, and let $\mu(X)$ be the cardinality of X. Then this condition is satisfied.

Given a probabilistic interpretation, we define a probability function, p, by:

$$p(\alpha) = \mu([\alpha])/\mu(I)$$

It is easy to see that p satisfies all the standard conditions for a probability function, such as:

$$0 \leq p(\alpha) \leq 1$$
$$\text{if } \alpha \models \beta \text{ then } p(\alpha) \leq p(\beta)$$
$$\text{if } \models \alpha \text{ then } p(\alpha) = 1$$
$$p(\alpha \vee \beta) = p(\alpha) + p(\beta) - p(\alpha \wedge \beta)$$

except, of course: $p(\neg\alpha) + p(\alpha) = 1$. Since we have $p(\alpha \wedge \neg\alpha) > 0$, and $p(\alpha \vee \neg\alpha) = 1$, it follows that $p(\alpha) + p(\neg\alpha) > 1$.

By the construction, we have, in fact, $p(\alpha) > 0$ for all α. It might be suggested that a person whose personal probability function gives nothing the value zero would have to be very stupid—or at least credulous. But since $p(\alpha)$ may be as small as one wishes, this hardly seems to follow. Moreover, giving nothing a zero probability signals an open-minded and undogmatic policy of belief. Arguably, this is the most rational policy.

Given a probability function, conditional probability can be defined in the usual way:

$$p(\alpha/\beta) = p(\alpha \wedge \beta)/p(\beta)$$

A singular advantage of this paraconsistent probability theory over standard accounts is that conditional probability is *always* defined, since the denominator is always non-zero.

[103]Again, many other paraconsistent logics could be used instead.

Perhaps the major application of probability theory is in framing an account of non-deductive inference. How, exactly, to do this is a moot question. But however one does it, a paraconsistent account of non-deductive inference can be framed in the same way, employing paraconsistent probability theory. For example, we may define the degree of (non-deductive) validity of the inference α/β to be $p(\beta/(\alpha \wedge \eta))$, where η is our background evidence. As one would expect, deductively valid inferences come out as having maximal degree of inductive validity.

To compute the degree of validity of an inference, so defined, we would often need to employ Bayes' Theorem. Let us look at the paraconsistent two-hypothesis version of this. Suppose that we have two hypotheses, h_1 and h_2, that are exclusive and exhaustive, in the sense that $\models h_1 \vee h_2$ and $\models \neg(h_1 \wedge h_2)$, and that we wish to compute the probability of h_1 on evidence, e, given the inverse probabilities of these hypotheses on the evidence (all relative to some background evidence, η, which we will ignore).

Note first that $p(h_1/e) = p(h_1 \wedge e)/p(e) = p(e/h_1).p(h_1)/p(e)$. It remains to compute $p(e)$. Since $h_1 \vee h_2$ entails $e \vee h_1 \vee h_2$ we have :

$$\begin{aligned} 1 = \quad p(e \vee h_1 \vee h_2) &= p(e) + p(h_1 \vee h_2) - p(e \wedge (h_1 \vee h_2)) \\ &= p(e) + 1 - p(e \wedge (h_1 \vee h_2)) \end{aligned}$$

Hence:

$$\begin{aligned} p(e) &= p(e \wedge (h_1 \vee h_2)) \\ &= p((e \wedge h_1) \vee (e \wedge h_2)) \\ &= p(e \wedge h_1) + p(e \wedge h_2) - p(e \wedge h_1 \wedge h_2) \\ &= p(h_1).p(e/h_1) + p(h_2).p(e/h_2) - p(h_1 \wedge h_2).p(e/(h_1 \wedge h_2)) \end{aligned}$$

Thus:

$$p(h_1/e) = \frac{p(e/h_1).p(h_1)}{p(h_1).p(e/h_1) + p(h_2).p(e/h_2) - p(h_1 \wedge h_2).p(e/(h_1 \wedge h_2))}$$

This is the paraconsistent version of Bayes' Theorem. In the classical case, the last term of the denominator is zero, since $\models \neg(h_1 \wedge h_2)$; but this is not so in the paraconsistent case. The theorem illustrates a general fact about paraconsistent probability theory: everything works as normal, except that we have to carry round some extra terms concerning the probabilities of certain contradictions which may be neglected in the classical case.

The extra complication may actually be a gain in some contexts. Let me mention one possible one; this concerns quantum mechanics. Quantum mechanics is known to suffer from various phenomena often called 'causal anomalies', a famous one of which is the two-slit experiment.[104] In this, a

[104] See, e.g., Haack [1974], ch. 8.

light is shone onto a screen through a mask with two slits. The intensity of light on any point on the screen is proportional to the probability that a photon hits it, σ, given that it goes through one slit, α, or goes through the other, β. Let us write $p(\alpha \vee \beta)$ as q. Then:

$$\begin{aligned} p(\sigma/(\alpha \vee \beta)) &= p(\sigma \wedge (\alpha \vee \beta))/p(\alpha \vee \beta) \\ &= p((\sigma \wedge \alpha) \vee (\sigma \wedge \beta))/q \\ &= p(\sigma \wedge \alpha)/q + p(\sigma \wedge \beta)/q - p(\sigma \wedge \alpha \wedge \beta)/q \end{aligned}$$

Classically, we know that $\neg(\alpha \wedge \beta)$, and so the last term may be ignored. For similar reasons, $q = p(\alpha \vee \beta) = p(\alpha) + p(\beta)$, and by symmetry we can arrange for $p(\alpha)$ and $p(\beta)$ to be equal. Hence:

$$\begin{aligned} p(\sigma/(\alpha \vee \beta)) &= p(\sigma \wedge \alpha)/2p(\alpha) + p(\sigma \wedge \beta)/2p(\beta) \\ &= \tfrac{1}{2}(p(\sigma/\alpha) + p(\sigma/\beta)) \end{aligned}$$

Thus, the intensity of light on the screen should be the average of the intensities of light going each slit independently (which can be determined by closing off the other). Exactly this is what is *not* found.

Standard quantum logic[105] avoids the result by rejecting the inference of distribution (i.e., the equivalence between $\sigma \wedge (\alpha \vee \beta)$ and $(\sigma \wedge \alpha) \vee (\sigma \wedge \beta)$, and so faulting the second line of the above proof. A paraconsistent solution is just to note that we cannot ignore the third term in the computation of $p(\sigma/(\alpha \vee \beta))$, even though we know that $\neg(\alpha \wedge \beta)$. In qualitative terms, what this means is that the photon has a non-zero probability of doing the impossible, and going through both slits simultaneously!

This application of paraconsistent probability theory to quantum mechanics is *highly* speculative. Whether it could be employed to resolve the other causal anomalies of quantum theory, let alone to predict the observations that are actually made, has not been investigated.[106]

7.6 The Classical Recapture

Most paraconsistent logicians have supposed that reasoning in accordance with classical logic is sometimes legitimate. Most, for example, have taken it that classical logic is perfectly acceptable in consistent situations. They have therefore proposed ways in which classical logic can be "recaptured" from a paraconsistent perspective.

The simplest such recapture occurs in non-adjunctive logics. As we noted in 4.2, single premise non-adjunctive reasoning is classical. Hence, classical

[105] See the article on this in the *Handbook*.

[106] For more on the above issues, including the effects of paraconsistent probability theory on confirmation theory, see Priest [1987], sect. 7.6, and Priest *et al.* [1989], pp. 376-9, 385-8.

reasoning can be regained simply by conjoining all premises. A different strategy is to employ a consistency operator, as is done in the da Costa logics C_i, for finite non-zero i. As we saw in 5.3, this can be employed to define a negation which behaves classically; hence classical reasoning can simply be interpreted within the system. This approach has problems for some applications, as we shall see when we come to look at set theory.

Yet another way to recapture classical reasoning, provided that a conditional operator is available, is to employ an absurdity constant, $\perp$, satisfying the condition $\models \perp \to \alpha$, for all α. Such a constant makes perfectly good sense paraconsistently. Algebraically, it corresponds to the minimal value of an algebra (which can usually be added if it is not present already). In truth-preservational terms, there are two ways of handling its semantics. One is to require that $\perp$ be untrue at every (world of every) evaluation. Its characteristic principle then holds vacuously. The other way (which may be preferable if one objects to vacuous reasoning) is simply to assign $\perp$ (at a world) the value of the (infinitary) conjunction of all other formulas (at that world). A bit of juggling then usually verifies the characteristic principle. (The definition itself guarantees it only when α does not contain $\perp$.)

Now let C be the set of all formulas of the form $(\alpha \wedge \neg\alpha) \to \perp$. Then an inference is classically valid iff it is enthymematically valid with C as the set of suppressed premises, in most paraconsistent logics. For if every member of C holds at (a world of) an interpretation, then the (world of the) interpretation is a classical one—or at least the trivial one—and hence if the premises of a classically valid inference are true at it, so is the conclusion. Thus, we have an enthymematic recapture.

Let us write $-\alpha$ for $\alpha \to \perp$. In classical (and intuitionist) logic, $-\alpha$ just is $\neg\alpha$. It might therefore be thought that provided a logic possesses $\perp$, we could simply interpret classical logic in it by identifying $\neg\alpha$ with $-\alpha$. This thought would be incorrect, though. In many paraconsistent logics, $-\alpha$ behaves quite differently from classical (and intuitionist) negation. What properties it has depends, of course, on the properties of $\to$. While it will always be the case that $\alpha, -\alpha \models \beta$, it will certainly not be true in general that $\models \alpha \vee -\alpha$, that $- - \alpha \models \alpha$, or even that $\alpha \models - - \alpha$. As an example of the last, consider an intensional interpretation for the logic H. (See 5.5.) Suppose that p is true at some normal world, w, but that at some non-normal world $p \to \perp$ is true (and $\perp$ is not). Then $(p \to \perp) \to \perp$ fails at w.[107]

A final, and much less brute-force, way of recapturing classical logic starts from the idea that consistency is the norm. It is implicit in the paraconsistent enterprise that inconsistency can be contained. Instead of spreading everywhere, inconsistencies can exist isolated, as do singularities in a field

[107]It might be thought that the existence of the explosive connective '$-$' would cause problems for certain paraconsistent applications; notably, for example, for set theory. This is not the case, however, as we will see.

(of the kind found in physics, not agriculture). This metaphor suggests that even if inconsistencies are present they will be relatively rare. If it is *true* inconsistencies we are talking about, these will be even rarer—something that the classical logician can readily agree with![108]

This suggests that consistency should be a default assumption, in the sense of non-monotonic logic. Many non-monotonic logics can be formulated by defining validity over some class of models, minimal with respect to violation of the default condition. In effect, we consider only those interpretations that are no more profligate in the relevant way than the information necessitates. In the case where it is consistency that is the default condition, we may define validity over models that are minimally inconsistent in some sense. I will illustrate, as usual, with respect to LP.[109]

Let $\mathcal{I} = \langle D, d \rangle$ be an LP interpretation. Let $\alpha \in \mathcal{I}!$ iff α is $Pd_1...d_n$, where P is an n-place predicate and $\langle d_1, ..., d_n \rangle \in E_P \cap A_P$ in $\mathcal{I}$. (Recall that I am using members of the domain as names for themselves.) $\mathcal{I}!$ is a measure of the inconsistency of $\mathcal{I}$. In particular, $\mathcal{I}$ is a classical interpretation iff $\mathcal{I}! = \phi$. If $\mathcal{I}_1$ and $\mathcal{I}_2$ are LP interpretations, I will write $\mathcal{I}_1 < \mathcal{I}_2$, and say that $\mathcal{I}_1$ is *more consistent than* $\mathcal{I}_2$, iff $\mathcal{I}_1! \subset \mathcal{I}_\in!$. (The containment here is proper.) $\mathcal{I}$ is a *minimally inconsistent (mi) model* of Σ iff $\mathcal{I}$ is a model of Σ iff $\mathcal{I}$ is a model of Σ and if $\mathcal{J} < \mathcal{I}, \mathcal{J}$ is not a model of Σ. Finally, α *is an mi consequence* of Σ ($\Sigma \models_m \alpha$) iff every mi model of Σ is a model of α.

As is to be expected, $\models_m$ is non-monotonic. For if p and q are atomic sentences, it is easy to check that $\{p, \neg p \vee q\} \models_m q$, but $\{\neg p, p, \neg p \vee q\} \not\models_m q$. Moreover, since all classical models (if there are any) are mi models, and all mi models are models, it follows that $\Sigma \models \alpha \Rightarrow \Sigma \models_m \alpha \Rightarrow \Sigma \models_C \alpha$. The implications are, in general, not reversible. For the first, note that $\{p, \neg p \vee q\} \not\models q$; for the second, note that $\{p, \neg p\} \not\models_m q$. But if Σ is classically consistent, its mi models are exactly its classical models, and hence we have $\Sigma \models_m \alpha \Leftrightarrow \Sigma \models_C \alpha$: classical recapture.

$\models_m$ has various other interesting properties. For example, it can be shown that if the LP consequences of some set is non-trivial, so are its mi consequences *Reassurance*. For details, see Priest [1991a].[110]

[108]Though this is not so obvious once one accepts dialetheism. For a defence of the view given dialetheism, see Priest [1987], sect. 8.4.

[109]Though the first paraconsistent logician to employ this strategy was Batens [1989], who employs a non-truth-functional logic. Batens also considers the dymanical aspects of such default reasoning.

[110]In that paper, in the definition of $<$, a clause stating that the domains of $\mathcal{I}_\infty$ and $\mathcal{I}_\in$ are the same is added. With this clause, the result concerning classical recapture is false (and that paper is mistaken). For example, if α is $\exists x Px \wedge \exists x \neg Px$, then $\langle D, d \rangle$ is an mi model, wher $D = \{a\}, E_P = A_P = \{a\}$, though this is not a classical model. (This was first noted by Diderik Batens, in correspondence.) As $<$ is defined here, $\{\forall x(Px \wedge \neg Px)\} \models_m \exists x \forall y x = y$, which may be thought to be counter-intuitive. But if $\forall x(Px \wedge \neg Px)$ is *all* the information we have, and inconsistencies are to be minimised, perhaps it *is* correct to infer that there is just one thing. Note that $\{\forall x(Px \wedge \neg Px), \exists x Qx \wedge \exists x \neg Qx\} \not\models_m \exists x \forall y x = y$. For $\langle d, d \rangle$ is an mi model of the premises, where $D = \{a, b\}, E_P = A_P =$

8 SEMANTICS AND SET THEORY

The previous part gestured in the direction of various applications of paraconsistent logic. I want, in the next two parts, to look at some other applications in greater detail. These concern theories of particular mathematical significance. In this part I will deal with semantics and set theory.

Semantic and set-theoretic notions appear to be governed by simple and apparently obvious principles. In semantics, these concern truth, T, satisfaction, S, and denotation, D, and are:

T-schema: $T\langle\alpha\rangle \leftrightarrow \alpha$
S-schema: $Sx\langle\beta\rangle \leftrightarrow \beta(y/x)$
D-schema: $D\langle t\rangle x \leftrightarrow x = t$

where α is any sentence, β is any formula with one free variable, y, and t is any closed term. Angle brackets indicate a name-forming device. In set theory the principle is the schema of set existence:

Comprehension Schema: $\exists x\forall y(y \in x \leftrightarrow \beta)$

where β is *any* formula not containing x. What the connective $\leftrightarrow$ is in the above schemas, we will have to come back to.

Despite the fact that these schemas appear to be obvious, they all give rise to contradictions, as is well known: the paradoxes of self-reference, such as (respectively) the Liar Paradox, the Heterological Paradox, Berry's Paradox and Russell's Paradox. The usual approaches to set theory and semantics restrict the principles in some way. Such approaches are all unsatisfactory in one way or another, though I shall not discuss this here.[111]

A paraconsistent approach can simply leave the principles as they are, and allow the contradictions to arise. They need do no damage, because the logic is not explosive. Even so, not all paraconsistent logics are suitable as the underlying logics of these theories. For a start, if the above schemas are formulated with the material $\equiv$ they give rise to a conjoined contradiction, so using a non-adjunctive logic (except Rescher and Manor's) explodes the theory.[112] And in the da Costa systems, C_i, for finite i, an operator behaving like classical negation, $\neg^*$ can be defined (see 5.3). The usual arguments establish contradictions of the form $\alpha \wedge \neg^*\alpha$, and so again

$\{a, b\}, E_Q = \{a\}, A_Q = \{b\}$. With the present definition, the proof of Reassurance for the first-order case, appropriately modified, still goes through.

[111] See, for example, Priest [1987], chs. 1, 2.

[112] Rescher and Brandom, [1980], p. 164, suggest splitting the biconditionals up into two non-conjoined conditionals.

the theories explode. Fortunately, there are other paraconsistent logics that will do the job.[113]

8.1 *Truth Theory in LP*

Let us start with the semantic case. I will deal with truth; similar remarks and constructions hold for the other semantic notions, but I will leave readers to ponder these for themselves. The first question we need to address is what connective it is that occurs in the biconditional of the T-schema. The first possibility is that it is a material biconditional, $\equiv$.[114]

Let us, then, suppose that we are dealing with the logic LP. We will need some machinery to handle self reference; a straightforward option is to let this be arithmetic. Hence, we suppose the language, L, to be that of first order arithmetic augmented by a one place predicate, T. To make things easy, we will assume that L has a function symbol for each primitive recursive function (and only those function symbols). Let T_0 be the LP theory in this language which comprises the truths of first order arithmetic plus the T-schema.

The assumption that T_0 contains all of arithmetic is obviously a very strong one, and means that the theory is not axiomatic. We could, instead, consider an axiomatic theory with some suitable fragment of arithmetic, but since a major part of our concern will be with what *cannot* be proved, it is useful to have the arithmetic part as strong as possible.

The first thing to note is that T_0 is inconsistent. Given the resources of arithmetic, for any formula, α, of one free variable, x, one can find, by the usual Gödel construction, a fixed-point formula, β, of the form $\alpha(x/\langle\beta\rangle)$.[115] Now, let α be $\neg Tx$ and let β be its fixed point. Then the T-schema gives us: $T\langle\beta\rangle \equiv \beta$, i.e., $T\langle\beta\rangle \equiv \neg T\langle\beta\rangle$. Unpacking the definition of $\equiv$, in terms of $\wedge$, $\vee$, and $\neg$ and fiddling, gives exactly $T\langle\beta\rangle \wedge \neg T\langle\beta\rangle$.[116]

Despite being inconsistent, T_0 is non-trivial. An easy way to see this is to observe, first of all, that if in any interpretation $\nu(\alpha) = \{1,0\}$ then $\nu(\alpha \equiv \beta) = \{1,0\}$. Hence, an LP model for T_0 can be obtained by letting the denotations of the arithmetic language be that of the standard interpretation of arithmetic—so that, in particular, the domain is N, the natural numbers; recall that classical interpretations are just special cases of LP

[113]There are paraconsistent set theories based on da Costa's C systems. (See, e.g., Arruda [1980], da Costa [1986].) In these theories, the schemas have to be constrained, as they are classically. This takes away much of the appeal of a paraconsistent approach.

[114]It is natural to suppose that it ought to be a detachable conditional. Goodship [1996] argues that it is only a material conditional. Whether or not this is the case, it is certainly interesting to explore the two possibilities.

[115]See, e.g., Priest [1987], sect. 3.5.

[116]It is worth noting that for the S-schema, the fixed point machinery is unnecessary for the demonstration of inconsistency. For let α be $\neg Sxx$. Then an instance of the S-schema is: $S\langle\alpha\rangle\langle\alpha\rangle \equiv \neg S\langle\alpha\rangle\langle\alpha\rangle$, and we can then proceed as before.

interpretations—and setting E_T and A_T, the extension and anti-extension of T, both to N. Call this interpretation $\mathcal{I}_0$. In $\mathcal{I}_0$ every sentence of the form $T\langle\alpha\rangle$ takes the value $\{1,0\}$, and so by the observation concerning $\equiv$, $\mathcal{I}_0$ is a model for the T-schema, and so of all of T_0. The same interpretation shows that if α is any arithmetic formula false in the standard model, $T_0 \not\models \alpha$.

T_0 is a relatively weak theory. In particular, it does not legitimate the two way rule of inference:

$$\frac{\alpha}{\overline{T\langle\alpha\rangle}}$$

(just consider the south-north inference in $\mathcal{I}_0$, where α is an arithmetic sentence false in the standard model).[117]

Let the theory obtained by replacing the material T-schema of T_0 with this rule be called T_1. T_1 is inconsistent. For choose an α of the form $\neg T\langle\alpha\rangle$. The law of excluded middle gives $T\langle\alpha\rangle \vee \neg T\langle\alpha\rangle$, i.e., $T\langle\alpha\rangle \vee \alpha$, which, applying the rule, gives $T\langle\alpha\rangle$ and α, i.e., $\neg T\langle\alpha\rangle$.

We can construct a model for T_1 as follows. If an interpretation assigns the standard denotations to all arithmetical language let us call it *arithmetical.* Any arithmetical interpretation is a model all of T_1 except, perhaps, the T-schema. Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be two arithmetical interpretations, with assignment functions ν_1 and ν_2. Define $\nu_1 \preceq \nu_2$ to mean that for all atomic sentences in the language, α:

$$\nu_1(\alpha) = t \Rightarrow \nu_2(\alpha) = t$$
$$\nu_1(\alpha) = f \Rightarrow \nu_2(\alpha) = f$$

If $\nu_1 \preceq \nu_2$ then this condition extends to all formulas of L. For suppose that $\nu_1 \preceq \nu_2$. If n is in the extension of T in $\mathcal{I}_2$ but not $\mathcal{I}_1$; then $\nu_2(Tn) = t$ or b, but $\nu_1(Tn) = f$, violating the condition. Similarly for anti-extensions. Hence, $\mathcal{I}_2 \leq \mathcal{I}_1$. By monotonicity, for all α, $\nu_2(\alpha) \subseteq \nu_1(\alpha)$. The conclusion follows. For suppose that $\nu_2(\alpha) \neq t$. Then α is false (i.e., b or f) in $\mathcal{I}_2$; hence α is false in $\mathcal{I}_1$, i.e., $\nu_1(\alpha) \neq t$. The argument for f is similar.

This result is, in fact, just another version of monotonicity; I will call it the *Monotonicity Lemma.*

Let $\mathcal{I}_0$ be any arithmetical interpretation, with evaluation function ν_0. We now define a transfinite sequence of arithmetical interpretations,

[117] Whether or not more follows with minimally inconsistent LP (see 7.6) is presently unknown. Another non-monotonic notion of inference also suggests itself here. According to this, the things that follow are the things that hold in all minimally inconsistent models where the arithmetic part is the standard model. Employing this would be appropriate if there were good reasons to believe that the only inconsistencies involve the truth predicate.

$\langle \mathcal{J}_i; i \in On \rangle$ (On is the class of ordinals). I will make the construction slightly more complex than necessary, for the benefits of the next section. It suffices to define the evaluation function ν_i of each interpretation. If $i > 0$ and n is not the code of a sentence, then $\nu_i(Tn) = \nu_0(Tn)$. We therefore need to consider only atomic formulas of the form $T\langle\alpha\rangle$. Let us say that α is *eventually t by k* iff $\exists i > 0 \forall j (i \leq j < k,\ \nu_j(\alpha) = t)$. Similarly for f. Then for $k \neq 0$:

$$\begin{array}{lll} \nu_k(T\langle\alpha\rangle) & = t & \text{if } \alpha \text{ is eventually } t \text{ by } k \\ & = f & \text{if } \alpha \text{ is eventually } f \text{ by } k \\ & = b & \text{otherwise} \end{array}$$

We can now establish that if $0 < i \leq k$ then $\nu_i \preceq \nu_k$. The proof is by transfinite induction. Suppose that the result holds for all $j < k$. We show it for k. Since the truth values of atomic formulas other than ones of the form $T\langle\alpha\rangle$ are constant, we need consider only these. So suppose that $\nu_i(T\langle\alpha\rangle) = t$. Then α is eventually t by i. In particular, for some $0 < j < i$, $\nu_j(\alpha) = t$. By induction hypothesis, for all l such that $j < l < k$, $\nu_j \preceq \nu_l$. Hence, by monotonicity $\nu_l(\alpha) = t$. Hence, α is eventually t by k, i.e., $\nu_k(T\langle\alpha\rangle) = t$. The case for f is similar.

What this lemma shows is that once $i > 0$, and increases, sentences of the form $T\langle\alpha\rangle$ can change their truth value at most once. If they ever attain a classical value, they keep it. Since there is only a countable number of sentences of this form, there must be an ordinal, l, by which all the formulas that change value have done so. Hence $\nu_l = \nu_{l+1}$. Call $\mathcal{J}_l$, $\mathcal{J}_*$; and its corresponding evaluation function ν^*. Then if $\nu^*(\alpha) = t$, $\nu^*(T\langle\alpha\rangle) = t$. Similarly for f and b. Hence $\nu^*(\alpha) = \nu^*(T\langle\alpha\rangle)$, and so $\mathcal{J}_*$ is a model of T_1. For the same reason, $\mathcal{J}_*$ also verifies the two-way rule:

$$\frac{\neg\alpha}{\overline{\neg T\langle\alpha\rangle}}$$

Yet the theory is not trivial: anything false in the standard model of arithmetic is untrue in $\mathcal{J}_*$, and so $T_1 \not\models \alpha$.

It is not difficult to see that the construction used to define $\mathcal{J}_*$ is, in fact, just a dualised form of Kripke's fixed point construction for a logic with truth value gaps using the strong Kleene three-valued logic.[118] (Provided we start with a suitable ground model, monotonicity is guaranteed from the beginning, and so we can just set $\nu_k(T\langle\alpha\rangle)$ to t (or f) if α takes the value t (or f) at some $i < k$.) Hence, if any sentence is *grounded* in Kripke's sense, it takes a classical value in $\mathcal{J}_*$. In particular, if α is any false grounded sentence, $T_1 \not\models \alpha$.

[118]See the article on Semantics and the Liar Paradox in this H*andbook*. One of the first people to realise that the construction could be dualised for this end was Dowden [1984].

8.2 Adding a Conditional

Although T_1 validates the two-way inferential T-schema, it does not validate the T-schema as formulated with a detachable conditional. This is for the simple reason that LP does not contain such a conditional. A natural thought is to augment the language with one to make this possible. Let the resulting language be $L_{\rightarrow}$. Not all conditionals are suitable here, however. This is due to Curry paradoxes. If the conditional satisfies the inference of contraction: $\alpha \rightarrow (\alpha \rightarrow \beta) \models \alpha \rightarrow \beta$, then the theory collapses into triviality. For consider the fixed-point formula, γ, of the form $T\langle\gamma\rangle \rightarrow \perp$ (or if $\perp$ is not present, just an arbitrary β). The T-schema gives: $T\langle\gamma\rangle \leftrightarrow (T\langle\gamma\rangle \rightarrow \perp)$. Contraction gives us: $T\langle\gamma\rangle \rightarrow \perp$ and then a couple of applications of *modus ponens* give $\perp$.[119]

This fact rules out the use of all the non-transitive logics we looked at (since they validate $\alpha \leftrightarrow (\alpha \rightarrow \beta) \models \beta$), all the da Costa logics and discussive logic (using discussive implication for the T-schema), since these validate contraction, and those relevance logics that validate contraction, such as R.[120] A relevant logic without contraction can be used for the purpose.

Let T_2 be as for T_0, except that the T-schema is formulated with $\rightarrow$, and the underlying logic is BX (see 5.5, 6.5). T_2 is inconsistent, since it is obviously stronger than T_1. But it can be shown to be non-trivial. If we try to generalise the proof for T_1 in simple ways, attempts are stymied by the failure of anything like monotonicity once $\rightarrow$ is involved. However, there is a way of building on the proof.[121] This requires us to move from objectual semantics to simple evaluational semantics. For the purpose of this section (and this one only), an atomic formula will be any of the usual kind *or* any one of the form $\alpha \rightarrow \beta$. Clearly, any sentence of the language can be built up from atomic formulas using $\wedge$, $\vee$, $\neg$, $\exists$ and $\forall$. Call an evaluation of atomic formulas, ν, *arithmetical* if it assigns to every identity its value in the standard model of arithmetic. Given an arithmetical evaluation, it is extended to an evaluation of all sentences by LP truth conditions, using substitutional quantification.

A quick induction shows that any arithmetical evaluation assigns t to all the arithmetic truths of the standard model (which do not contain $\rightarrow$ or T), and f to all the falsehoods. Moreover, for this notion of valuation, we do have the Monotonicity Lemma. Finally, given any such evaluation, we

[119] An argument of this kind first appeared in Curry [1942]. Different versions that employ close relatives of contraction, such as $\vdash (\alpha \wedge (\alpha \rightarrow \beta)) \rightarrow \beta$ (but *not* $\alpha \wedge (\alpha \rightarrow \beta) \vdash \beta$) can also be found in the literature. See, e.g., Meyer *et al.* [1979].

[120] For good measure, it also rules out using Rescher and Manor's non-adjunctive approach. Using this, every consistent sentence would follow, since if β is consistent, so is $\alpha \leftrightarrow (\alpha \rightarrow \beta)$.

[121] The following is taken from Priest [1991b], which simply modifies Brady's proof for set theory in [1989].

can construct a fixed point, ν^*, such that $\nu^*(\alpha) = \nu^*(T\langle\alpha\rangle)$, as in 8.1. The construction is the same, except that in the definition of ν_k, we set Ts to t if s is a (closed) term that evaluates to the code of α, and α is eventually t by k. Similarly for f. (The values of atoms of the form $\alpha \to \beta$ do not change in the process.)

An induction shows that if $\mu \preceq \nu$ then for all i, $\mu_i \preceq \nu_i$. Suppose the result for all $i < k$. We show it for k. We need consider only those atomic formulas of the form Ts where s evaluates to the code of sentence α. $\mu_k(Ts) = t$ iff α is eventually t by k, for μ. By induction hypothesis, this implies that α is eventually t by k for ν. Hence, $\nu_k(Ts) = t$, as required. The case for f is similar. From this result it obviously follows that if $\mu \preceq \nu$ then $\mu^* \preceq \nu^*$.

Let $\Rightarrow$ be the conditional connective of $RM3$ (see 5.4, identifying $+1$, 0, and -1 with t, b, and f, respectively). This also plays a role in the proof. Its relevant property is that if $\mu \preceq \nu$ then if α and β are formulas of $L_\to$ and $\mu(\alpha \Rightarrow \beta) = t$, $\nu(\alpha \Rightarrow \beta) = t$. For if $\mu(\alpha \Rightarrow \beta) = t$ then $\mu(\alpha) = f$ or $\mu(\beta) = t$. By monotonicity $\nu(\alpha) = f$ or $\nu(\beta) = t$. Hence, $\nu(\alpha \Rightarrow \beta) = t$.

Let ν_0 be the arithmetical interpretation that assigns every sentence of the form Ts the value b. We now define a transfinite sequence of arithmetic valuations, $\langle \nu_i; i \in On\rangle$, as follows. (I write $(\nu_j)^*$ as ν_j^*.) For $k \neq 0$:

$$\begin{aligned}
\nu_k(\alpha \to \beta) &= t \qquad \text{if } \forall j < k,\ \nu_j^*(\alpha \Rightarrow \beta) = t\\
&= f \qquad \text{if } \exists\, j < k,\ \nu_j^*(\alpha \Rightarrow \beta) = f\\
&= b \qquad \text{otherwise}
\end{aligned}$$

And where α is of the form Ts, where s is any closed term which evaluates to the code of a sentence:

$$\begin{aligned}
\nu_k(\alpha) &= t \qquad \text{if } \exists i \forall j (i \leq j < k, \nu_j^*(\alpha) = t)\\
&= f \qquad \text{if } \exists i \forall j (i \leq j < k, \nu_j^*(\alpha) = f)\\
&= b \qquad \text{otherwise}
\end{aligned}$$

We can now establish that if $i \leq k$ then $\nu_i \preceq \nu_k$. The proof is by transfinite induction. Suppose that the result holds for all $j < k$. We need to consider cases where a formula is of the form $\alpha \to \beta$ or Ts, where s is a term that evaluates to the code of a sentence. Take them in that order.

Suppose that $\nu_i(\alpha \to \beta) = t$. Then $\nu_0^*(\alpha \Rightarrow \beta) = t$. By induction hypothesis, for $0 < j < k$, $\nu_0 \preceq \nu_j$. Thus, $\nu_0^* \preceq \nu_j^*$. Hence, $\nu_j^*(\alpha \Rightarrow \beta) = t$, by the observation concerning $\Rightarrow$. Thus, $\nu_k(\alpha \to \beta) = t$, as required. The case for f is trivial.

For the other case, suppose that $\nu_i(\alpha) = t$. Then $\exists j < i$, $\nu_j^*(\alpha) = t$. By induction hypothesis, if $j \leq l < k$, $\nu_j \preceq \nu_l$, and hence $\nu_j^* \preceq \nu_l^*$. By monotonicity, $\nu_l^*(\alpha) = t$. Thus, $\nu_k(\alpha) = t$. The case for f is similar.

What this lemma shows, as before, is that we must eventually reach an l such that $\nu_l = \nu_{l+1}$. Let this evaluation be $\vec{\nu}$. Then $\vec{\nu}$ is a model of all the extensional arithmetic apparatus. It also models the T-schema. For if $i < l$, $\nu_i^*(\alpha) = \nu_i^*(T\langle\alpha\rangle)$, and so $\nu_i^*(\alpha \Leftrightarrow T\langle\alpha\rangle) = t$ or b, and $\vec{\nu}(\alpha \leftrightarrow T\langle\alpha\rangle) = t$ or b. (For the same reason, $\vec{\nu}$ models the contraposed form: $\neg\alpha \leftrightarrow \neg T\langle\alpha\rangle$. Since $T\langle\neg\alpha\rangle \leftrightarrow \neg\alpha$ is an instance of the T-schema, it also models $T\langle\neg\alpha\rangle \leftrightarrow \neg T\langle\alpha\rangle$.)

It remains to check that $\vec{\nu}$ models the axioms and respects the rules of inference of BX. This requires no little checking. Most of it is routine. Here, for example, is one of the harder propositional axioms: $((\alpha \to \beta) \wedge (\alpha \to \gamma)) \to (\alpha \to ((\beta \wedge \gamma))$. Let the antecedent be φ, and the consequent be ψ. Then $\vec{\nu}(\varphi \to \psi) = t$ or b iff for no $i < l$, $\nu_i^*(\varphi \Rightarrow \psi) = f$. Now, suppose that $\nu_i^*(\varphi \Rightarrow \psi) = f$. Then one of:

$\nu_i^*(\varphi) = t$ and $(\nu_i^*(\psi) = b$ or $\nu_i^*(\psi) = f)$
$\nu_i^*(\varphi) = b$ and $\nu_i^*(\psi) = f$

In the first case, $\nu_i(\alpha \to \beta) = \nu_i(\alpha \to \gamma) = t$. But then for all $j < i$, $\nu_j^*(\alpha \Rightarrow \beta) = \nu_j^*(\alpha \Rightarrow \gamma) = t$, in which case $\nu_j^*(\alpha \Rightarrow (\beta \wedge \gamma)) = t$, and so $\nu_i(\alpha \to (\beta \wedge \gamma)) = t$, which is impossible. In the second case, $\nu_i(\alpha \to \beta) = t$ or b, and $\nu_i(\alpha \to \gamma)) = t$ or b. But then for all $j < i$, $\nu_j^*(\alpha \Rightarrow \beta) = t$ or b, and $\nu_j^*(\alpha \Rightarrow \gamma) = t$ or b, in which case $\nu_j^*(\alpha \Rightarrow (\beta \wedge \gamma)) = t$ or b, and so $\nu_i(\alpha \to ((\beta \wedge \gamma)) = t$ or b, which is also impossible.

For further details, see Brady [1989].[122] The construction shows that T_1 is non-trivial, since if α is any arithmetic sentence false in the standard model $\vec{\nu}(\alpha) = f$. (Indeed, as with the previous construction, which is incorporated in this, if α is any false grounded sentence, the same is true.)

8.3 Advantages of a Paraconsistent Approach

What we have seen is that it is possible to have a theory containing all the machinery of arithmetic, plus a truth predicate which satisfies the T-schema for every sentence of the language—whether this is formulated as a material biconditional, a two-way rule of inference, or a detachable bi-conditional. It is inconsistent, but non-trivial; in fact, the inconsistencies do not spread

[122]Brady shows that the construction verifies propositional logics that are a good deal stronger than BX. His treatment of identiy is different, though. To verify the substitutivity rule of 7.1, it suffices to show that if $t_1 = t_2$ holds in an interpretation then $\alpha(x/t_1)$ and $\alpha(x/t_2)$ have the same truth value. A quick induction shows that if this is true for atomic α it is true for all α. Hence, we need consider only these. Next, show by induction that if this holds for ν it holds for all evaluations in the construction of ν^*, and so of ν^* itself. Finally, we show by induction that it holds for every ν_i in the hierarchy, and hence for $\vec{\nu}$.

into the arithmetic machinery.[123] Thus, it is possible to have a workable, if inconsistent, theory which respects the central intuition about truth.

It is not my aim here to discuss the shortcomings of other standard approaches to the theory of truth,[124] but none can match this. All restrict the T-schema in one way or another. The one that comes closest to having the full T-schema is Kripke's account of truth, which at least has it in the form of a two way rule of inference. However, this account has the singular misfortune of being self-referentially inconsistent. According to this account, if α is the Liar sentence it is *neither* true nor false, and so not true, but the theory pronounces $\neg T\langle\alpha\rangle$ itself neither true nor false. According to T_2, α is *both* true and false (i.e., has a true negation), and this is exactly what it proves: $T\langle\alpha\rangle \wedge \neg T\langle\alpha\rangle$ entails $T\langle\alpha\rangle \wedge T\langle\neg\alpha\rangle$. It might also show that α is *not* true (and so not both true and false). But paraconsistency shows you exactly how to live with this kind of contradiction.

This is not unconnected with the matter of "strengthened" paradoxes. If someone holds the Liar sentence to be neither true nor false, one can invite them to consider the sentence, β, 'This sentence is not true' (as opposed to false). Whether β is true, false or neither, a contradiction arises. It is sometimes suggested that a paraconsistent account of truth falls to the same problem, since β can have no consistent truth-value on this account either. It should be clear that this argument is just an *ignoratio*. A paraconsistent account does not *require* it to have a consistent truth-value. In fact, according to T_2, $T\langle\neg\alpha\rangle \leftrightarrow \neg T\langle\alpha\rangle$; if this is right, there is no distinction between the standard Liar and the "strengthened Liar" at all.[125]

Let me finish with a word of caution. We can construct non-trivial theories which incorporate the S-schema of satisfaction and the D-schema of denotation, in exactly the same way as we did the T-schema. If, however, we try to add descriptions to a theory with self-reference and the D-schema, trouble does arise.

Suppose that we have a description operator, ε, satisfying the Hilbertian principle: $\exists x\alpha \vdash \alpha(x/\varepsilon x\alpha)$. If t is any closed term, $t = t$, and so by the D-schema $D\langle t\rangle t$, and $\exists x D\langle t\rangle x$. Thus, by the description principle, $D\langle t\rangle \varepsilon x D\langle t\rangle x$, whence, by the D-schema again:

$$t = \varepsilon x D\langle t\rangle x$$

[123] Nor does the T-schema have to be taken as axiomatic. One can give truth conditions for atomic sentences and then prove the T-schema in the usual Tarskian fashion. See Priest [1987], ch. 7.

[124] For this, see Priest [1987], ch. 2.

[125] The advantages of a paraconsitent account of truth rub off onto any account of modal (deontic, doxastic, etc.) operators that treats them as predicates. For all such theories are just sub-theories of the theory of truth. See Priest [1991b]. We will have an application of this concerning provability in 9.6.

Now in arithmetic, just as for any formula, α, with one free variable, x, we can find a sentence, β, of the form $\alpha(x/\langle\beta\rangle)$, so, for any term, t, with one free variable, x, we can find a closed term, s, such that s is $t(x/\langle s\rangle)$. If f is any one place function symbol, apply this fact to the term $f\varepsilon yDxy$, to obtain an s such that:

$$s = f\varepsilon yD\langle s\rangle y$$

Since $s = \varepsilon yD\langle s\rangle y$, it follows that $s = fs$: any function has a fixed point. This shows that the semantic machinery does have purely arithmetic consequences. In particular, for example, $\exists x\, x = x + 1$. Arithmetic statements like this can be kept under control, as we will see later in the next part, but worse is to come.

Let f be the parity function, i.e.:

$$\begin{aligned} fx &= 0 \quad \text{if } x \text{ is odd} \\ &= 1 \quad \text{if } x \text{ is even} \end{aligned}$$

We have $fs = 0 \vee fs = 1$. In the first case $s = fs = 0$, which is even, and so $fs = 1$. Thus, $0 = 1$. Similarly in the second case. This is unacceptable, even for someone who supposes that there are *some* inconsistent numbers.

Where to point the finger of suspicion is obvious enough. As we saw, the D-schema entails $\exists xD\langle t\rangle x$, for any closed term, t; and there is no reason why someone who subscribes to a paraconsistent account of semantic notions must believe that every term has a denotation: in particular, in the vernacular, 's' is 'a number that is 1 if it is even and 0 if it is odd', which would certainly seem to have no denotation. This suggests that the D-schema should be subjected to the condition that $\exists xD\langle t\rangle x$ in some suitable way. The behaviour of resulting theories is a particularly interesting unsolved problem.[126]

8.4 Set Theory in LP

Let us now turn to the second theory that we will look at, set theory. This is a theory of sets governed by the full Comprehension schema. This schema is structurally very similar to the T-schema, and many of the considerations of previous subsections carry over to set theory in a straightforward manner. The major element of novelty concerns the other axiom, the Extensionality axiom.

Let us start with set theory in LP. The language here contains just the predicates $=$ and $\in$, and the axioms are:

[126]For a further discussion of all of these issues, see Priest [1997a].

$$\exists x \forall y(y \in x \equiv \alpha)$$
$$\forall x(x \in y \equiv x \in z) \supset y = z$$

where x does not occur free in α. Call this theory S_0. S_0 is inconsistent. For putting $y \notin y$ for α, and instantiating the quantifier we get: $\forall y(y \in r \equiv y \notin y)$, whence $r \in r \equiv r \notin r$. Cashing out $\supset$ in terms of $\neg$ and $\vee$ gives $r \in r \wedge r \notin r$.

In constructing models of S_0, the following observation (due to Restall [1992]) is a useful one. First some definitions. Given two vectors of LP values, $(g_m; m \in D)$, $(h_m; m \in D)$, the first ***subsumes*** the second iff for all $m \in D$, $g_m \supseteq h_m$. Now consider a matrix of such values $(e_{m,n}; m, n \in D)$. This is said to *cover* the vector $(g_m; m \in D)$ iff for some $n \in D$, the vector $(e_{m,n}; m \in D)$ subsumes it. A vector indexed by D is *classical* iff all its members are t or f. (Recall that we are writing $\{1\}$, $\{1,0\}$, $\{0\}$ as t, b f, respectively.)

Now the observation. Consider an LP interpretation, $\langle D, d\rangle$, and the matrix $(e_{m,n}; m, n \in D)$, where $e_{m,n} = \nu(m \in n)$. If this covers every classical vector indexed by D it verifies the Comprehension principle. For let α be any formula not containing x, and consider the vector $(\nu(\alpha(y/m)); m \in D)$. This certainly subsumes some classical vector; choose one such, and let this be subsumed by $(e_{m,n}; m \in D)$. Now consider any formula of the form $m \in n \equiv \alpha(y/m)$. Where the two sides differ in value, one of them has the value b. Hence, the value of the biconditional is either t or b. Thus the same is true of $\forall y(y \in n \equiv \alpha)$, and $\exists x \forall y(y \in x \equiv \alpha)$.

Using this fact, it is easy to construct models for S_0. Consider an LP interpretation, $\langle D, d\rangle$, where $D = \{m, n\}$, and $e_{m,n}$ is given by the following matrix:

$\in$	m	n
m	b	t
n	b	t

Each column is the membership vector of the appropriate member of D; and since that of m subsumes every classical vector indexed by D, this verifies the Comprehension axiom. In the Extensionality axiom, if y and z are the same, the axiom is obviously true. If they are distinct, one is n and the other is m, and for each x, the value of $x \in n \equiv x \in m$ is b. Hence, $\forall x(x \in n \equiv x \in m)$ has the value b and Extensionality is verified. In this model, $m \notin n$ and $n \notin n$ have the value f, as, therefore do $\exists y\, y \notin n$ and $\forall x \exists y\, y \notin x$. Hence, S_0 is non-trivial.

A characterisation of what can be proved in S_0 (and of what its minimally inconsistent consequences are) is still an open question. There are, however, certainly theorems of Zermelo Fraenkel set theory, ZF, that are not provable in S_0. For example, in ZF there is provably no universal set: $ZF \vdash \forall x \exists y\, y \notin x$. But this is not a consequence of S_0, as we have just seen.[127]

The simple model of S_0 that we have just used to prove non-triviality is obviously pathological in some sense. An interesting question is what the "intended" interpretations of S_0 are like. Whilst unable to give an answer to this, I note that for any classical model of ZF, $\mathcal{M} = \langle D, d \rangle$, there is a model of S_0 which has $\mathcal{M}$ as a substructure. Let a be some new object, let $\mathcal{M}^+ = \langle D^+, d^+ \rangle$, where $D^+ = D \cup \{a\}$, d^+ is the same as d, except that for every $c \in D^+$, the value of $c \in a$ is b; for every $c \in D$, the value of $a \in c$ is f; the value of $a = a$ is t; and for every $c \in D$, the value of $a = c$ is f. $\mathcal{M}$ is clearly a substructure of $\mathcal{M}^+$. The membership vector of a subsumes every classical vector, and hence $\mathcal{M}$ is a model of Comprehension.

It remains to verify Extensionality: $\forall x(x \in m \equiv x \in n) \supset m = n$. If m and n are the same in $\mathcal{M}^+$, then the consequent is true, as is the conditional. So suppose that they are distinct. If they are both in D, then, by extensionality in $\mathcal{M}$, there is some $c \in D$ such that $c \in m$ is t and $c \in n$ is f, or vice versa. Whichever of these is the case, $c \in m \equiv c \in n$ is f, as is $\forall x(x \in m \equiv x \in n)$. Hence the conditional is t. Finally, suppose that $m \in D$ and n is a (or vice versa, which is similar). Then if $c \in D^+$, every sentence of the form $c \in n$ is b. Hence, every sentence of the form $c \in m \equiv c \in n$ is b, as therefore is $\forall x(x \in n \equiv x \in m)$. Hence, the conditional is true.

8.5 Brady's Non-triviality Proof

As a working set theory, S_0 is rather weak. Since the Comprehension axiom is only a material one, we cannot infer that something is in a set from the fact that it satisfies its defining condition, and vice versa. This suggests strengthening the principle to a two-way rule of inference, as we did for truth theory. This, in turn, requires the addition of set abstracts to the language. So let us enrich the language with terms of the form $\{x; \alpha\}$ for any variable, x, and formula, α; and trade in the Comprehension principle of S_0 for the two-way rule:

$$\frac{x \in \{y; \alpha\}}{\overline{\alpha(y/x)}}$$

Call this theory S_1. S_1 is inconsistent. For let r be $\{x; x \notin x\}$. Then:

$$\frac{r \in r}{\overline{r \notin r}}$$

The law of excluded middle then quickly gives us $r \in r \wedge r \notin r$.

[127]For this, and some further observations in this direction, see Restall [1992].

The non-triviality of S_1 is presently an open question. But even though it is probably non-trivial, as a working set theory, it is still rather weak. This is because we have no useful way of establishing that two sets are identical. Even if we can show that $\forall x(\alpha \equiv \beta)$, and so that $\forall x(x \in \{x;\alpha\} \equiv x \in \{x;\beta\})$, we cannot infer that $\{x;\alpha\} = \{x;\beta\}$ since Extensionality does not support a detachable inference.

We might hope to circumvent this problem by trading in the Extensionality principle for the corresponding rule:

$$\frac{\forall x(\alpha \equiv \beta)}{\{x;\alpha\} = \{x;\beta\}}$$

But if we do this, trouble arises.[128] For let r be as before. Then since $r \in r$ must take the value b in any interpretation, we have, for any α, $\forall x(\alpha \equiv r \in r)$, and so $\{x;\alpha\} = \{x; r \in r\}$. Thus, for any α and β, $\{x;\alpha\} = \{x;\beta\}$; which is rather too much.

The problem arises because the Extensionality rule of inference allows us to move from an equivalence that does not guarantee substitution ($\alpha \equiv \beta, \beta \equiv \gamma \not\models_{LP} \alpha \equiv \gamma$) to one that does (identity). This suggests formulating Extensionality itself with a connective that legitimises substitution. So let us add a detachable connective to the language, $\rightarrow$, and formulate Extensionality as:

$$\frac{\forall x(\alpha \leftrightarrow \beta)}{\{x;\alpha\} = \{x;\beta\}}$$

The trouble then disappears.

And now that we have a detachable conditional connective at our disposal, it is natural to formulate the Comprehension principle as a detachable biconditional, as follows:

$$\forall y(y \in \{x;\alpha\} \leftrightarrow \alpha(x/y))$$

We have to be careful about the conditional connective here. As with truth, any conditional connective that satisfies contraction would give rise to triviality. For let c be $\{x \in x \rightarrow \bot\}$. Then an instance of Comprehension is $y \in c \leftrightarrow (y \in y \rightarrow \bot)$. Instantiating with c, we get $c \in c \leftrightarrow (c \in c \rightarrow \bot)$, and we can then proceed, as with truth, to obtain $\bot$. Even if the logic does not contain contraction, Curry-style paradoxes may still be forthcoming. For example, if we drop the contraction axiom from the relevant logic R

[128] There are other cases where the full Comprehension principle by itself is alright, but throwing in extensionality causes problems; for example, set theory based on Łukasiewicz' continuum-valued logic. See White [1979].

and add the law of excluded middle, the Comprehension principle still gives triviality.[129] Again however, a relevant logic without contraction will do the job.

Consider the set theory with Extensionality and Comprehension formulated as just described, and based on the underlying logic BX (with free variables, so that these may occur in the schematic letters of Extensionality and Comprehension). Call this S_2. The first thing to note about S_2 is that identity can be defined in it, in Russellian fashion. Writing $x = y$ for $\forall z(x \in z \leftrightarrow y \in z)$, $x = x$ follows. Substituting $\{w; \alpha\}$ for z, and using the Comprehension principle gives $\alpha(w/x) \leftrightarrow \alpha(w/y)$. Hence, we need no longer assume that $=$ is part of the language.

Since the Comprehension principle of S_2 gives the two-way deduction version of S_1, S_2 is inconsistent. It is also demonstrably non-trivial, as shown by Brady [1989].[130] To prove this, we repeat the proof for T_2 of 8.2 with three modifications. The first, a minor one, is that we add two propositional constants t and f to the language; their truth values are always what the letters suggest. (This is necessary to kick-start the generation of the fixed point into motion. In the case of truth, this was done by the arithmetic sentences.) More substantially, in constructing ν^* we replace the clause for T by:

$$\begin{array}{lll} \nu_k(s \in \{x; \alpha\}) & = t & \text{if } \alpha(x/s) \text{ is eventually } t \text{ by } k \\ & = f & \text{if } \alpha(x/s) \text{ is eventually } f \text{ by } k \\ & = b & \text{otherwise} \end{array}$$

where s is any closed term, and α contains at most x free. The final modification is that in extending evaluations to all formulas, we use substitutional quantification with respect to the closed set abstracts.

Now, $\vec{\nu}$ verifies all the theorems of S_2, in the sense that if α is any closed substitution instance of a theorem, it receives the value t or b in $\vec{\nu}$. This is shown by an induction on the length of proofs. That the logical axioms have this property, and the logical rules of inference preserve this property, is shown as in 8.2. This leaves the set theoretic ones.

Given the construction of $\vec{\nu}$, it is not difficult to see that it verifies the Comprehension principle. It is not at all obvious that Extensionality preserves verification. What needs to be shown is that if $\forall x(\alpha \leftrightarrow \beta)$ is verified, so is anything of the form $a \in c \leftrightarrow b \in c$, where a is $\{x; \alpha\}$ and b is $\{x; \beta\}$. Let c be $\{y; \gamma\}$. Then, given Comprehension, what needs to be shown is that $\gamma(y/a) \leftrightarrow \gamma(y/b)$ is verified. If this can be shown for atomic γ, the result will follow by induction. Given the premise of the inference and Comprehension, it is true if γ is of the form $d \in y$. If it is of the form $y \in d$, where

[129] See Slaney [1989]. Other classical principles are also known to give rise to triviality in conjunction with the Comprehension schema. See Bunder [1986].

[130] A modification of the proof shows that the theory based on the logic B is, in fact, consistent. See Brady [1983].

d is $\{z;\delta\}$, we need to show that $\delta(z/a) \leftrightarrow \delta(z/b)$ is verified. We obviously have a regress. In fact, the regression grounds out in a suitable way in the construction of $\vec{\nu}$. For details, see Brady [1989].[131]

The non-triviality of S_2 is established since there are many sentences that are not verified by $\vec{\nu}$. It is easy to check, for example, that any sentence of the form $c \in \{x; f\}$ takes the value f, as, therefore, does the formula $\forall x \exists y\, y \in x$.

A notable feature of Brady's proof is the following. As formulated, the Comprehension principle entails: $\exists y \forall x(x \in y \leftrightarrow \alpha)$, where y does not occur in α. (The y in question is $\{x;\alpha\}$ and so cannot be a subformula of α.) If we relax the restriction, we get an absolutely unrestricted version of the principle. Brady's proof can be extended to verify this version too, by adding a fixed point operator to the language, and treating it suitably. Again, for details, see Brady [1989].

Finally, it is worth observing that the T-schema is interpretable in S_2. If α is any closed formula, let us write $\langle\alpha\rangle$ for $\{z;\alpha\}$, where z is some fixed variable. Define Tx to be $a \in x$, where a is any fixed term. Then $T\langle\alpha\rangle = a \in \{z;\alpha\} \leftrightarrow \alpha$. Moreover, the absolutely unrestricted Comprehension principle gives us fixed points of the kind required for self-reference. Let α be any formula of one free variable, x. By the principle, there is a set, s, such that $\forall x(x \in s \leftrightarrow \alpha(x/\{z; a \in s\}))$. It follows that $a \in s \leftrightarrow \alpha(x/\{z; a \in s\})$. Thus, if β is $a \in s$, we have $\beta \leftrightarrow \alpha(x/\langle\beta\rangle)$. S_2 (with the absolutely unrestricted Comprehension principle) therefore gives us a demonstrably non-trivial *joint* theory of truth, sethood and self-reference.

8.6 *Paraconsistent Set Theory*

Despite the strong structural similarities between semantics and set theory, there is an important historical difference. Set theory is a well developed mathematical theory in a way that semantics is not. In the case of set theory, it is therefore natural to ask how a paraconsistent theory such as S_2 relates to this development.

To answer this question (at least to the extent that the answer is known), it will be useful to divide set theory into three parts. The first comprises that basic set-theory which all branches of mathematics use as a tool. The second is transfinite set theory, as it can be established in ZF. The third

[131] Brady's treatment of identity is slightly different from the one given here. He defines $x = y$ as $\forall z(z \in x \leftrightarrow z \in y)$. Given Comprehension, this delivers the version of Extensionality used here straight away. What is lost is the substitution principle $x = y, \alpha(w/x) \vdash \alpha(w/y)$. Given the Comprehension principle, this can be reduced to $x = y, x \in z \vdash y \in z$ (which follows from our definition of identity). Brady takes something stronger than this as his substitutivity axiom: $\vdash (x = y \wedge z = z) \rightarrow (x \in z \leftrightarrow y \in z)$. Hence, his construction certainly verifies the weaker principle. It is worth noting that the construction does not validate the simpler $\vdash x = y \rightarrow (x \in z \leftrightarrow y \in z)$, which, in any case, is known to be a Destroyer of Relevance. See Routley [1980b], sect. 7.

concerns results about sets, like Russell's set and the universal set, that do not exist in ZF. Let us take these matters in turn.

S_2 is able to provide for virtually all of bread-and-butter set theory (Boolean operations on sets, power sets, products, functions, operations on functions, etc.), and so provide for the needs of working mathematics.[132] For example, if we define the Boolean operators, $x \cap y$, $x \cup y$ and $\bar{x}$ as $\{z; z \in x \wedge z \in y\}$, $\{z; z \in x \vee z \in y\}$ and $\{z; z \notin x\}$, respectively, and $x \subseteq y$ as $\forall z(z \in x \rightarrow z \in y)$, then we can establish the usual facts concerning these notions. Some care needs to be taken over defining a universal set, U, and empty set, ϕ, though. If we define ϕ, as $\{x; x \neq x\}$, we cannot show that for all y, $\phi \subseteq y$, since the underlying logic is relevant and cannot prove $x \neq x \rightarrow \alpha$ for arbitrary α. (Dually for U.) If we define ϕ as $\{x; \forall z\ x \in z\}$, this problem is solved, since $\forall z\ x \in z \rightarrow x \in y$. (Dually for U.)

The reason for the qualification 'virtually' in the first sentence of the last paragraph, is as follows. The sets, as structured by union, intersection and complementation, are not a Boolean algebra, but a De Morgan algebra with maximum and minimum elements. Though we can show that $\forall y\ y \notin x \cap \bar{x}$, we cannot show that $x \cap \bar{x} \subseteq \phi$, since, relevantly, $(\alpha \wedge \neg\alpha) \rightarrow \beta$ fails. (Dually for U.) There are, in a sense, more than one universal and empty sets. Moreover, this is essential. If we had $x \cap \bar{x} \subseteq \phi$ then, taking $\{z; \alpha\}$ for x, we get $(\alpha \wedge \neg\alpha) \rightarrow \forall y\ z \in y$. Now take $\{z; \beta\}$ for y, and we get $(\alpha \wedge \neg\alpha) \rightarrow \beta$; paraconsistency fails. In fact, Dunn [1988] shows that if the principles that there is a unique universal set, and a unique empty set, are added to any set theory such as S_2, full classical logic falls out.

Turning to the second area, the question of how much of the usual transfinite set theory can be established in S_2 is one to which the answer is currently unknown. What can be said is that the *standard* proofs of a number of results break down. This is particularly the case for results that are proved by *reductio*, such as Cantor's Theorem. Where α is an assumption made for the purpose of *reductio*, we may well be able to establish that $(\alpha \wedge \beta) \rightarrow (\gamma \wedge \neg\gamma)$, for some γ, where β is the conjunction of other facts appealed to in deducing the contradiction (such as instances of the Comprehension principle). But contraposing and detaching will give us only $\neg\alpha \vee \neg\beta$, and we can get no further.[133]

Lastly, the third area: reasoning in S_2, one can prove various results about sets that are impossible in ZF. For example, as usual, let $\{x\}$ be $\{y; y = x\}$, $\{x, y\}$ be $\{x\} \cup \{y\}$ and $\bigcup x$ be $\{z; \exists y \in x,\ z \in y\}$. $r = \{x; x \notin x\}$, and we know that $r \in r$ and $r \notin r$. Then:[134]

(1) If $x \in r$ then $\{x\} \in r$. For $\{x\} \in \{x\}$ or $\{x\} \notin \{x\}$. In the first case,

[132]Much of this is spelled out in Routley [1980b], sect. 8.

[133]Interesting enough, however, it is possible to prove a version of the Axiom of Choice using the completely unrestricted version of the Comprehension principle. See Routley [1980b], sect. 8.

[134]The following is taken from Arruda and Batens [1982].

$\{x\} = x$, and so $\{x\} \in r$. In the second case, $\{x\} \in r$ by definition.

(2) If $x, y \in r$ then $\{x, y\} \in r$. For $\{x, y\} \in \{x, y\}$ or $\{x, y\} \notin \{x, y\}$. In the first case, $\{x, y\} = x$ or $\{x, y\} = y$, and so $\{x, y\} \in r$. In the second case, $\{x, y\} \in r$ by definition.

(3) $\{\{x, r\}\} \in r$. For $\{\{x, r\}\} \in \{\{x, r\}\}$ or $\{\{x, r\}\} \notin \{\{x, r\}\}$. In the first case, $\{\{x, r\}\} = \{x, r\}$, hence, $x = \{x, r\} = r$. But then $x, r \in r$ so $\{x, r\} \in r$, by (2), and, $\{\{x, r\}\} \in r$, by (1). In the second case, $\{\{x, r\}\} \in r$ by definition.

(4) $\forall x\ x \in \bigcup r$. For suppose that $\{x, r\} \in \{x, r\}$. Then $\{x, r\} = x$ or $\{x, r\} = r$. In the first case, $\{x\} = \{\{x, r\}\}$, so $\{x\} \in r$, by (3). In the second, $\{x, r\} \in r$. In either case $x \in \bigcup r$. Suppose, on the other hand, that $\{x, r\} \notin \{x, r\}$. Then $\{x, r\} \in r$, by definition, and so $x \in \bigcup r$.

That $\bigcup r$ is universal, is hardly a profound result. But it at least illustrates the fact that there are possibilities which transcend ZF.

Let me end this section with a speculative comment on what all this shows. The discussion of this section, and especially the part concerning the non-Boolean properties of sets in S_2, shows that it is impossible to recapture standard set theory in its entirety in this theory. Sets are extensional entities *par excellence*; using an intensional connective in their identity conditions is bound to gum up the works. In fact, it seems to me that the most plausible way of viewing S_2 is as a theory of properties, where intensional identity conditions are entirely appropriate. But what you call these entities does not really matter here. The important fact is that they are not the sets of standard modern mathematical practice.

If we want a theory of such entities, the appropriate identity conditions must employ $\equiv$, and this means that we are back with the proof-theoretically weak S_0 (or S_1). Since this does not contain ZF, how should someone who subscribes to a paraconsistent theory of such sets view modern mathematical practice?

One answer is as follows. The standard model of ZF is the cumulative hierarchy. As we saw in 8.2, there are models of S_0 which contain this hierarchy. We may thus take it that the intended interpretation of S_0 is a model of this kind (or if there are more than one, that they are all models of this kind). The cumulative hierarchy is therefore a (consistent) fragment of the set-theoretic universe, and modern set theory provides a description of it. There is, however, more to the universe than this fragment. A classical logician may well agree with that claim. For example, they may think that there are also non-well-founded sets. The paraconsistent logician agrees with this: after all, r is not well-founded; but they will think that sets outside the hierarchy may have even more remarkable properties: some of them are inconsistent.

9 ARITHMETIC AND ITS METATHEORY

In this part I want to look at the application of paraconsistent logic to another important mathematical theory: arithmetic. The situation concerning arithmetic is rather different from that concerning set theory and semantics. There are no apparently obvious and intrinsically arithmetical principles that give rise to contradiction, in the way that the Comprehension principle and the T-schema do—or if there are, this fact has not yet been discovered. In the first instance, the paraconsistent interest in arithmetic arises because there is a class of inconsistent models of arithmetic. (It might be more accurate to say 'models of inconsistent arithmetic'.) It may be supposed that these models are pathological in some sense.[135] I will come back to this matter later. But even if it is so, the models nevertheless have an interesting and important *mathematical* structure, as do the classical non-standard models of arithmetic—which are, in fact, just a special case, as we will see. And just as one does not have to be an intuitionist to find intuitionistic structures of intrinsic mathematical interest, so one does not have to be a dialetheist for the same to be true of inconsistent structures. One thing this part illustrates, therefore, is the existence of a new branch of mathematics which concerns the investigation of just such structures.[136]

The existence of inconsistent models of arithmetic bears, as might be expected, on the limitative theorems of Metamathematics. And whatever the status of the inconsistent models themselves, many have held that these theorems have important philosophical implications. This part will also look at the connection between the inconsistent models and the limitative theorems, and I will comment on the significance of this for the philosophical implications of Gödel's incompleteness theorem.

9.1 The Collapsing Lemma

Let us start with a theorem about LP on which much of the following depends: the Collapsing Lemma.[137]

Let $\mathcal{I} = \langle D, d\rangle$ be any interpretation for LP. Let $\sim$ be any equivalence relation on D, that is also a congruence relation on the denotations of the function symbols in the language (i.e., if g is such a denotation, and $d_i \sim e_i$ for all $1 \leq i \leq n$, then $g(d_1, ..., d_n) \sim g(e_1, ..., e_n)$). If $d \in D$ let $[d]$ be the

[135] Though this claim has certainly been queried. See Priest [1994].

[136] On this, see further, Mortensen [1995]. Perhaps surprisingly, the first person to investigate an inconsistent arithmetic was Nelson [1959], who gave a realisability-style semantics for the language of arithmetic, according to which the set of formulas realised was inconsistent (and closed under a logic somewhat weaker than intuitionist logic).

[137] The theorem works equally well for FDE, but we will be concerned primarily with models of theories that contain the law of excluded middle, and so where there are no truth-value gaps.

equivalence class of d under $\sim$. Define an interpretation, $\mathcal{I}_\sim = \langle D_\sim, d_\sim \rangle$, to be called the *collapsed interpretation*, where $D_\sim = \{[d]; d \in D\}$; if c is a constant, $d_\sim(c) = [d(c)]$; if f is an n-place function symbol:

$$d_\sim(f)([d_1], ..., [d_n]) = [d(f)(d_1, ..., d_n)]$$

(this is well defined, since $\sim$ is a congruence relation); and if P is an n-place predicate, its extension and anti-extension in $\mathcal{I}_\sim$, $E_P^\sim$ and $A_P^\sim$, are defined by:

$$\begin{array}{l} \langle [d_1], ..., [d_n] \rangle \in E_P^\sim \text{ iff for all } 1 \leq i \leq n, \exists e_i \sim d_i, \langle e_1, ..., e_n \rangle \in E_P \\ \langle [d_1], ..., [d_n] \rangle \in A_P^\sim \text{ iff for all } 1 \leq i \leq n, \exists e_i \sim d_i, \langle e_1, ..., e_n \rangle \in A_P \end{array}$$

where E_P and A_P are the extension and anti-extension of P in $\mathcal{I}$. It is easy to check that $E_=^\sim$ is $\{\langle [d], [d] \rangle ; d \in D\}$, as required for an LP interpretation.

The collapsed interpretation, in effect, identifies all members of an equivalence class to produce a composite individual that has the properties of all of its members. It may, of course, be inconsistent, even if its members are not.

A swift induction confirms that for any closed term, t, $d_\sim(t) = [d(t)]$. Hence:

$$\begin{array}{rcl} 1 \in \nu(Pt_1...t_n) & \Rightarrow & \langle d(t_1), ..., d(t_n) \rangle \in E_P \\ & \Rightarrow & \langle [d(t_1)], ..., [d(t_n)] \rangle \in E_P^\sim \\ & \Rightarrow & \langle d_\sim(t_1), ..., d_\sim(t_n) \rangle \in E_P^\sim \\ & \Rightarrow & 1 \in \nu_\sim(Pt_1...t_n) \end{array}$$

Similarly for 0 and anti-extensions. Monotonicity then entails that for any formula, α, $\nu(\alpha) \subseteq \nu_\sim(\alpha)$. This is the *Collapsing Lemma.*[138]

The Collapsing Lemma assures us that if an interpretation is a model of some set of sentences, then any interpretation obtained by collapsing it will also be a model. This gives us an important way of constructing inconsistent models. In particular, if the language contains no function symbols, and $\mathcal{I}$ is a model of some set of sentences, then, by appropriate choice of equivalence relation, we can collapse it down to a model of *any* smaller size. Thus we have a very strong downward Löwenheim-Skolem Theorem: If a theory in a language without function symbols has a model, it has a model of all smaller cardinalities.

I note that, since monotonicity holds for second order LP (section 7.2), the Collapsing Lemma extends to second order LP. Details are left as an exercise.

[138]The result is proved in Priest [1991a]. A similar result was proved by Dunn [1979].

9.2 Collapsed Models of Arithmetic

From now on, let L be the standard language of first-order arithmetic: one constant, $\mathbf{0}$, function symbols for successor, addition and multiplication, $'$, $+$, and $\times$, respectively, and one predicate symbol, $=$. If $\mathcal{I}$ is any interpretation, let $Th(\mathcal{I})$ (the *theory* of $\mathcal{I}$) be the set of all sentences true in $\mathcal{I}$. Let $\mathcal{N}$ be the standard model of arithmetic, and $A = Th(\mathcal{N})$. Let $\mathcal{M} = \langle M, d\rangle$ be any classical model of A—which is just special cases of an LP model. (As is well known, there are many of these other than $\mathcal{N}$.[139]) I will refer to the denotations of $'$, $+$, and $\times$ as the arithmetic operations of $\mathcal{M}$, and since no confusion is likely, use the same signs for them.[140]

Let $\sim$ be an equivalence relation on M, that is also a congruence relation with respect to the interpretations of the function symbols. Then we may construct the collapsed interpretation, $\mathcal{M}_\sim$. By the Collapsing Lemma, $\mathcal{M}_\sim$ is a model of A. Provided that $\sim$ is not the trivial equivalence relation, that relates each thing only to itself, then $\mathcal{M}_\sim$ will model inconsistencies. For suppose that $\sim$, relates the distinct members of M, n and m, then in $\mathcal{M}_\sim$, $[n] = [m]$ and so $\langle [n], [m]\rangle$ is in the extension of $=$. But since $n \neq m$ in $\mathcal{M}$, $\langle [n], [m]\rangle$ is in the anti-extension too. Thus, $\exists x(x = x \wedge x \neq x)$ holds in $\mathcal{M}_\sim$.

As an illustration of constructing an inconsistent model of A using the Collapsing Lemma, suppose that we partition M into $n+1$ successive blocks, $C_0, ..., C_{n+1}$, such that if $x, z \in C_i$ and $x < y < z$ then $y \in C_i$. And suppose that for $0 < i \leq n+1$, C_i is closed under the arithmetic operations of $\mathcal{M}$. (The existence of such a partition follows from a standard result in the study of classical models of arithmetic. See Kaye [1991], sect. 6.1.) Let $1 \leq k \in C_0 \cup C_1$ and define $x \sim y$ as:

$(x, y \in C_0$ and $x = y)$ or
for some $0 < i \leq n+1$, $x, y \in C_i$ and $x = y \bmod k$

where '$x = y \bmod k$' means that for some $j \in M$, $x + j \times k = y$, in $\mathcal{M}$.

It is not difficult to check that $\sim$ is an equivalence relation on M, and, moreover, that it is a congruence relation on the arithmetic operations of $\mathcal{M}$. Hence, we may use it to give a collapsed model. In this, C_0 collapses into an initial tail of numbers, and each C_i $(0 < i \leq n+1)$ collapses into a block of period k. For example, if $\mathcal{M}$ is the standard model, $n = 1$ and $C_0 = \phi$, the collapsed model is a simple cycle of period k. The successor function in the model may be depicted as follows:

[139] See, e.g., Kaye [1991].
[140] For a more detailed discussion of the material in this section, see Priest [1997a].

$$\begin{array}{ccccccc} 0 & \rightarrow & 1 & \rightarrow & \ldots & \rightarrow & i \\ \uparrow & & & & & & \downarrow \\ k-1 & \leftarrow & & \ldots & & \leftarrow & i+1 \end{array}$$

I will call such models *cycle models*. They were, in fact, the first inconsistent models to be discovered.[141] If $\mathcal{M}$ is any model, $n = 1$, and $k = 1$, we have a tail isomorphic to C_0, and then a degenerate single-point cycle. In particular, if $\mathcal{M}$ is a non-standard model and C_0 comprises the standard numbers, we have the natural numbers with a "point at infinity", Ω:

$$0 \rightarrow 1 \rightarrow \ldots \quad \Omega \hookleftarrow$$

9.3 Inconsistent Models of Arithmetic

Now that we have seen the existence of inconsistent models of arithmetic, let us look at their general structure.

Take any LP model of arithmetic, $\mathcal{M} = \langle M, d\rangle$. I will call the denotations of the numerals *regular* numbers. Let $x \leq y$ be defined in the usual way, as $\exists z\ x + z = y$. It is easy to check that $\leq$ is transitive. For if $i \leq j \leq k$ then for some x, y, $i + x = j$ and $j + y = k$. Hence $(i + x) + y = k$. But $(i + x) + y = i + (x + y)$ (since it is a model of arithmetic). The result follows.

If $i \in M$, let $N(i)$ (the *nucleus* of i) be $\{x \in M; i \leq x \leq i\}$. In a classical model, $N(i) = \{i\}$, but this need not be the case in an inconsistent model. For example, in a cycle model the members of the cycle constitute a nucleus. If $j \in N(i)$ then $N(i) = N(j)$. For if $x \in N(j)$ then $i \leq j \leq x \leq j \leq i$, so $x \in N(i)$, and similarly in the other direction. Thus, every member of a nucleus defines the same nucleus.

Now, if N_1 and N_2 are nuclei, define $N_1 \preceq N_2$ to mean that for some (or all, it makes no difference) $i \in N_1$ and $j \in N_2$, $i \leq j$. It is not difficult to check that $\preceq$ is a partial ordering. Moreover, since for any i and j, $i \leq j$ or $j \leq i$, it is a linear ordering. The least member of the ordering is $N(0)$. If $N(1)$ is distinct from this, it is the next (since for any x, $x \leq 0 \vee x \geq 1$), and so on for all regular numbers.

Say that $i \in M$ has period $p \in M$ iff $i + p = i$. In a classical model every number has period 0 and only 0. But again, this need not be the case in an inconsistent model, as the cycle models demonstrate. If $i \leq j$ and i has period p so does j. For $j = i + x$, so $p + j = p + i + x = i + x = j$. In particular, if p is a period of some member of a nucleus, it is a period of

[141] This was by Meyer [1978]. Things are spelled out in Meyer and Mortensen [1984]. The idea of collapsing non-standard classical models is to be found in Mortensen [1987]. Different structures can be collapsed to provide inconsistent models of other kinds of number, e.g., real numbers. See Mortensen [1995].

every member. We may thus say that p is a period of the nucleus itself. It also follows that if $N_1 \preceq N_2$ and p is a period of N_1 it is a period of N_2.

If a nucleus has a regular non-zero period, m, then it must have a minimum (in the usual sense) non-zero period, since the sequence $0, 1, 2, ..., m$ is finite. If $N_1 \preceq N_2$ and N_1 has minimum regular non-zero period, p, then p is a period of N_2. Moreover, the minimum non-zero period of N_2, q, must be a divisor (in the usual sense) of p. For suppose that $q < p$, and that q is not a divisor of p. For some $0 < k < q$, p is some finite multiple of q plus k. So if $x \in N_2$, $x = x + q = x + p + ... + p + k$. Hence $x = x + k$, i.e., k is a period of N_2, which is impossible.

If a nucleus has period $p \geq 1$, I will call it *proper.* Every proper nucleus is closed under successors. For suppose that $j \in N$ with period p. Then $j \leq j' \leq j + p = j$. Hence, $j' \in N$. In an inconsistent model, a number may have more than one predecessor, i.e., there may be more than one x such that $x' = j$. (Although $x' = y' \supset x = y$ holds in the model, we cannot necessarily detach to obtain $x = y$.)[142] But if j is in a proper nucleus, N, it has a unique predecessor in N. For let the period of N be q'. Then $(j + q)' = j + q' = j$. Hence, $j + q$ is a predecessor of j; and $j \leq j + q' = j$. Hence, $j + q \in N$. Next, suppose that x and y are in the nucleus, and that $x' = y' = j$. We have that $x \leq y \vee y \leq x$. Suppose, without loss of generality, the first disjunct. Then for some z, $x + z = y$; so $j + z = j$, and z is a period of the nucleus. But then $x = x + z = y$. I will write the unique predecessor of j in the nucleus as $'j$.

Now let N be any proper nucleus, and $i \in N$. Consider the sequence $...,'' i,' i, i,$
$i', i''....$ Call this the *chromosome* of i. Note that if $i,\ j \in N$, the chromosomes of i and j are identical or disjoint. For if they have a common member, z, then all the finite successors of z are identical, as are all its finite predecessors (in N). Thus they are identical. Now consider the chromosome of i, and suppose that two members are identical. There must be members where the successor distance between them is a minimum. Let these be j and $j'^{\cdots\prime}$ where there are n primes. Then $j = j + n$, and n is a period of the nucleus—in fact, its minimum non-zero period—and the chromosome of every member of the nucleus is a successor cycle of period n.

Hence, any proper nucleus is a collection of chromosomes, all of which are either successor cycles of the same finite period, or are sequences isomorphic to the integers (positive and negative). Both sorts are possible in an inconsistent model. Just consider the collapse of a non-standard model, of the kind given in the last section, by an equivalence relation which leaves all the standard numbers alone and identifies all the others modulo p. If p is standard, the non-standard numbers collapse into a successor cycle; if it

[142] In fact, it is not difficult to show that there is at most one number with multiple predecessors; and this can have only two.

is non-standard, the nucleus generated is of the other kind.

To summarise so far, the general structure of a model is a liner sequence of nuclei. There are three segments (any of which may be empty). The first contains only improper nuclei. The second contains proper nuclei with linear chromosomes. The final segment contains proper nuclei with cyclical chromosomes of finite period. A period of any nucleus is a period of any subsequent nucleus; and in particular, if a nucleus in the third segment has minimum non-zero period, p, the minimum non-zero period of any subsequent nucleus is a divisor of p. Thus, we might depict the general structure of a model as follows (where $m+1$ is a multiple of $n+1$):

$$
\begin{array}{cccccc}
0,1,\ldots & \boxed{\begin{array}{c} \ldots a \to a' \ldots \\ \ldots b \to b' \ldots \\ \vdots \end{array}} & \ldots & \boxed{\begin{array}{ccccc} d_0 \ldots d_i & & e_0 \ldots e_i & & \\ \uparrow \quad \downarrow & & \uparrow \quad \downarrow & & \ldots \\ d_m \ldots d_i' & & e_m \ldots e_i' & & \end{array}} & \ldots \\
 & & \ldots & \boxed{\begin{array}{ccccc} f_0 \ldots f_i & & g_0 \ldots g_i & & \\ \uparrow \quad \downarrow & & \uparrow \quad \downarrow & & \ldots \\ f_n \ldots f_i' & & g_n \ldots g_i' & & \end{array}} & \ldots
\end{array}
$$

Another obvious question is what possible orderings the proper nuclei can have. For a start, they can have the order-type of any ordinal. To prove this, one establishes by transfinite induction that for any ordinal, α, there is a classical model of arithmetic in which the non-standard numbers can be partitioned into a collection of disjoint blocks with order-type α, closed under arithmetic operations. One then collapses this interpretation in such a way that each block collapses into a nucleus.

The proper nuclei need not be discretely ordered. They can also have the order-type of the rationals. To prove this, one considers a classical non-standard model of arithmetic, where the order-type of the non-standard numbers is that of the rationals. It is possible to show that these can be partitioned into a collection of disjoint blocks, closed under arithmetic operations, which themselves have the order-type of the rationals. One can then collapse this model in such a way that each of the blocks collapses into a proper nucleus, giving the result. This proof can be extended to show that any order-type that can be embedded in the rationals in a certain way, can also be the order-type of the proper nuclei. This includes ω^* (the reverse of ω) and $\omega^* + \omega$, but *not* $\omega + \omega^*$. For details of all this, see Priest [1997b].

What other linear order-types proper nuclei may or may not have, is still an open question.

9.4 Finite Models of Arithmetic

First-order arithmetic has many classical nonstandard models, but none of these is finite. One of the intriguing features of *LP* is that it permits finite models of arithmetic, e.g., the cycle models. For these, a complete characterisation is known.

Placing the constraint of finitude on the results of the previous section, we can infer as follows. The sequence of improper nuclei is either empty or is composed of the singletons of $0, 1, ..., n$, for some finite n. There must be a finite collection of proper nuclei, $N_1 \preceq ... \preceq N_m$; each N_i must comprise a finite collection of successor cycles of some minimum non-zero finite period, p_i. And if $1 \leq i \leq j \leq m$, p_j must be a divisor of p_i.[143]

Moreover, there are models of any structure of this form. To show this, we can generalise the construction of 9.2. Take any non-standard classical model of arithmetic. This can be partitioned into the finite collection of blocks:

$$C_0, C_{1_0}, ..., C_{1_{k(1)}}, ...C_{i_0}, ..., C_{i_{k(i)}}, ..., C_{m_0}, ..., C_{m_{k(m)}}$$

where C_0 is either empty or is of the form $\{0, ..., n\}$, each subsequent block is closed under arithmetic operations, and there are $k(i)$ successor cycles in N_i. We now define a relation, $x \sim y$, as follows:

$(x, y \in C_0$ and $x = y)$ or
for some $1 \leq i \leq m$:
(for some $0 < j < k(i)$, $x, y \in C_{i_j}$, and $x = y \bmod p_i$) or
($x, y \in C_{i_0} \cup C_{i_{k(i)}}$ and $x = y \bmod p_i$)

One can check that $\sim$ is an equivalence relation, and also that it is a congruence relation on the arithmetic operations. Hence we can construct the collapsed model. $\sim$ leaves all members of C_0 alone. For every i it collapses every C_{i_j} into a successor cycle of period p_i, and it identifies the blocks C_{i_0} and $C_{i_{k(i)}}$. Thus, the sequence $C_{i_0}, ...C_{i_{k(i)}}$ collapses into a nucleus of size $k(i)$. The collapsed model therefore has exactly the required structure.[144]

There are many interesting questions about inconsistent models, even the finite ones, whose answer is not known. For example: how many models of each structure are there? (The behaviour of the successor function in a model does not determine the behavior of addition and multiplication, except in the tail.) Perhaps the most important question is as follows. Not all inconsistent model of arithmetic are collapses of classical models. Let $\mathcal{M}$ be any model of arithmetic; if $\mathcal{M}'$ is obtained from $\mathcal{M}$ by adding extra pairs to the anti-extension of $=$, call $\mathcal{M}'$ an *extension* in $\mathcal{M}$. If $\mathcal{M}'$ is an extension of $\mathcal{M}$, monotonicity ensures that it is a model of arithmetic. Now, consider the extension of the standard model obtained by adding $\langle 0, 0 \rangle$ to the anti-extension of $=$. This is not a collapsed model, since, if it were, 0 would have to have been identified with some $x > 0$. But then 1 would have

[143]It is also possible to show that each nucleus is closed under addition and multiplication.

[144]For further details, see Priest [1997a].

been identified with $x' > 1$. Hence, $\mathbf{0}' \neq \mathbf{0}'$ would also be true in the model, which it is not. Maybe, however, each inconsistent model is the extension of a collapsed classical model. If this conjecture is correct, collapsed models can be investigated via an analysis of the classical models of arithmetic and their congruence relations.

9.5 The Limitative Theorems of Metamathematics

Let us now turn to the limitative theorems of Metamathematics in the context of LP. These are the theorems of Löwenheim-Skolem, Church, Tarski and Gödel. I will take them in that order.[145] In what follows, let P be the set of theorems of classical Peano Arithmetic, and let Q be any non-trivial theory that contains P.

According to the classical Löwenheim-Skolem, Q has models of every infinite cardinality but has no finite models. Moving to LP changes the situation somewhat. Q still has a model of every infinite cardinality.[146] But it has models of finite size too: any inconsistent model may be collapsed to a finite model merely by identifying all numbers greater than some cut-off.[147]

The situation with second order P is again different in LP. Classically, this is known to be categorical, having the standard model as its only interpretation. But as I noted in 9.1, the Collapsing Lemma holds for second order LP. Hence, second order P is not categorical in LP. For example, it has finite models.

Turning to Church's theorem, this says, classically, that Q is undecidable. In LP, extensions of Q may be decidable. For example, let $\mathcal{M}$ be any finite model of A $(= Th(\mathcal{N}))$, and let Q be $Th(\mathcal{M})$. Then Q is a theory that contains P. Yet Q is decidable, as is the theory of any finite interpretation. In the language of $\mathcal{M}$ there is only a finite number of atomic sentences; their truth values can be listed. The truth values of truth functions of these can be computed according to (LP) truth tables, and the truth values of quantified sentences can be computed, since $\exists x\alpha$ has the same truth value

[145]For a statement of these in the classical context, see Boolos and Jeffrey [1974]. This section expands on the appendix of Priest [1994].

[146]The standard classical proof of this adds a new set of constants, $\{c_i, i \in I\}$, to the language, and all sentences of the form $c_i \neq c_j$, $i \neq j$, to Q. It then uses the compactness theorem. Things are more complex in LP, since the fact that $c_i \neq c_j$ holds in an interpretation does not mean that the denotations of these constants are distinct. After extending the language, we observe that $c_i = c_j$ cannot be proved. We then construct a prime theory in the manner of 4.3, keeping things of this form *out*. This is then used to define an appropriate interpretation.

[147]Let us say that M is an ***exact model*** of a theory iff the truths of M are exactly the members of the theory. Classically, for complete theories, there is no difference between modelling and exact modelling. The situation for LP is more complex. It can be shown that if Q has an infinite exact model it has exact models of every greater cardinality. On the other hand, if Q has a finite model, $\mathcal{M}$, in which every number is denoted by a numeral, $\mathcal{M}$ can be shown to be the only exact model of Q (up to isomorphism).

as the disjunction of all formulas of the form $\alpha(x/a)$, where a is in the domain of $\mathcal{M}$; dually for $\forall$.

Tarski's Theorem: this says that Q cannot contain its own truth predicate, in the sense that even if Q is a theory in an extended language, there is no formula, β, of one free variable, x, such that $\beta(x/\langle\alpha\rangle) \equiv \alpha \in Q$, for all closed formulas, α, of the language. This, too, fails for LP. Let $\mathcal{M}$ be any (classical) model of P, let $\mathcal{M}'$ be any finite collapse of $\mathcal{M}$, and let Q be $Th(\mathcal{M}')$. By the Collapsing Lemma, Q contains P. Since Q is decidable, it is representable in (classical) P by a formula, β, of one free variable, x. That is, we have :

If $\alpha \in Q$ then $\beta(x/\langle\alpha\rangle) \in P$
If $\alpha \notin Q$ then $\neg\beta(x/\langle\alpha\rangle) \in P$

By the Collapsing Lemma, 'P' may be replaced by 'Q'. If $\alpha \in Q$, $\beta(x/\langle\alpha\rangle) \in Q$, and so $\beta(x/\langle\alpha\rangle) \equiv \alpha \in Q$ (since $\gamma, \delta \models_{LP} \gamma \equiv \delta$); and if $\alpha \notin Q$, $\neg\beta(x/\langle\alpha\rangle) \in Q$, and so $\beta(x/\langle\alpha\rangle) \equiv \alpha \in Q$ (since $\neg\alpha \in Q$ and $\neg\gamma, \neg\delta \models_{LP} \gamma \equiv \delta$).

There is no guarantee that α and $\beta(x/\langle\alpha\rangle)$ have the same truth value in $\mathcal{M}'$. In particular, then, β may not satisfy the T-schema in the form of a two-way rule of inference. So it might be said that β is not really a truth predicate. Whether or not this is so, we have already seen that there are Qs where there is a predicate satisfying this condition (though this has to be added to the language of arithmetic): the theory T_1 of section 8.1.[148]
Finally, let us turn to Gödel's undecidability theorems. A statement of the first of these is that if Q is axiomatisable then there are sentences true in the standard model that are not in Q. It is clear that this may fail in LP. Let $\mathcal{M}$ be any finite model of arithmetic. Then if Q is $Th(\mathcal{M})$, Q contains all of the sentences true in the standard model of arithmetic, but is decidable, as we have noted, and hence axiomatisable (by Craig's Theorem).

It is worth asking what happens to the "undecidable" Gödel sentence in such a theory. Let β be any formula that represents Q in Q. (There are such formulas, as we just saw.) Then a Gödel sentence is one, α, of the form $\neg\beta(x/\langle\alpha\rangle)$. If $\alpha \in Q$ then $\neg\beta(x/\langle\alpha\rangle) \in Q$, but $\beta(x/\langle\alpha\rangle) \in Q$ by representability. If $\alpha \notin Q$ then $\neg\beta(x/\langle\alpha\rangle) \in Q$ by representability, i.e., $\alpha \in Q$, so $\beta(x/\langle\alpha\rangle) \in Q$ by representability. In either case, then,

[148] The construction of 8.1 can be applied to any model of arithmetic—not just the standard model—as the ground model. However, if we apply it to a finite model care needs to be exercised. The construction will not work as given, since different formulas may be coded by the same number in the model, which renders the definition of the sequence of interpretations illicit. We can switch to evaluational semantics, as in 8.2, though the construction then no longer validates the substitutivity of identicals. Alternatively, we can refrain from using numbers as names, but just augment the language with names for all sentences.

$\alpha \wedge \neg\alpha \in Q$.

Gödel's second undecidability theorem says that the statement that canonically asserts the consistency of Q is not in Q; this statement is usually taken to be $\neg\beta(x/\langle\alpha_0\rangle)$, where α_0 is $\mathbf{0} = \mathbf{0}'$, and β is the canonical proof predicate of Q. This also fails in LP.[149] Let Q be as in the previous two paragraphs. Then Q is not consistent. However, it is still the case that $\alpha_0 \notin Q$ (provided that the collapse is not the trivial one). Consider the relationship: n is (the code of) a proof of formula (with code) m in Q. Since this is recursive, it is represented in A by a formula $Prov(x, y)$. If α is provable in Q then for some n, $Prov(\mathbf{n}, \langle\alpha\rangle) \in A$ (where $\mathbf{n}$ is the numeral for n); thus, $\exists x Prov(x, \langle\alpha\rangle) \in A$ and so Q. If α is not provable in Q then for all n, $\neg Prov(\mathbf{n}, \langle\alpha\rangle) \in A$; thus, $\forall x \neg Prov(x, \langle\alpha\rangle) \in A$ (since A is ω-complete) and $\neg\exists x Prov(x, \langle\alpha\rangle) \in A$ and so Q. Thus, $\exists x Prov(x, y)$ represents Q in Q. In particular, since $\alpha_0 \notin Q$, $\neg\exists x Prov(x, \langle\alpha_0\rangle) \in Q$, as required.

9.6 The Philosophical Significance of Gödel's Theorem

People have tried to make all sorts of philosophical capital out of the negative results provided by the limitative theorems of classical Metamathematics. As we have seen, all of these, save the Löwenheim-Skolem Theorem, fail for arithmetic based on a paraconsistent logic. Setting this theorem aside, then, nothing can be inferred from these negative results unless one has reason to rule out paraconsistent theories. At the very least, this adds a whole new dimension to the debates in question.

This is not the place to discuss all the philosophical issues that arise in this context, but let me say a little more about one of the theorems by way of illustration. Doubtless, the incompleteness result that has provoked most philosophical rumination is Gödel's first incompleteness theorem: usually in a form such as: for any axiomatic theory of arithmetic (with sufficient strength, etc.), which we can recognise to be sound, there will be an arithmetic truth—viz., its Gödel sentence—not provable in it, but which we can establish as true.[150] This is just false, paraconsistently. If the theory is inconsistent, the Gödel sentence may well be provable in the theory, as we have seen.

An obvious thought at this point is that if we can recognise the theory to be sound then it can hardly be inconsistent. But unless one closes the question prematurely, by a refusal to consider the paraconsistent possibility, this is by no means obvious. What *is* obvious to anyone familiar with the subject, is that at the heart of Gödel's theorem, is a paradox. The paradox concerns the sentence, γ, 'This sentence is not provable', where 'provable'

[149]It is worth noting that there are *consistent* arithmetics based on some relevant logics, notably R, for which the statement of consistency is in the theory. See Meyer [1978].

[150]For example, the theorem is stated in this form in Dummett [1963]; it also drives Lucas' notorious [1961], though it is less clearly stated there.

is not to be understood to mean being the theorem of some axiom system or other, but as meaning 'demonstrated to be true'. If γ is provable, then it is true and so not provable. Thus we have proved γ. It is therefore true, and so unprovable. Contradiction. The argument can be formalised with one predicate, B, satisfying the conditions:

$$\vdash B\langle\alpha\rangle \rightarrow \alpha$$
$$\text{If} \vdash \alpha \text{ then} \vdash B\langle\alpha\rangle$$

for all closed α—including sentences containing B. For if γ is of the form $\neg B\langle\gamma\rangle$, then, by the first, $\vdash B\langle\gamma\rangle \rightarrow \neg B\langle\gamma\rangle$, and so $\vdash \neg B\langle\gamma\rangle$, i.e., $\vdash \gamma$. Hence, $\vdash B\langle\gamma\rangle$, by the second.

And we *do* recognise these principles to be sound. Whatever is provable is true, by definition; and demonstrating α shows that α is provable, and so counts as a demonstration of this fact.[151]

B is a predicate of numbers, but we do not have to assume that B is definable in terms of $'$, $+$ and $\times$ using truth functions and quantifiers. The argument could be formalised in a language with B as primitive. As we saw in the previous part in connection with truth, it is quite possible to have an inconsistent theory with a predicate of this kind, where the sentences definable in terms of $'$, $+$ and $\times$ using truth functions and quantifiers behave quite consistently.

Of course, if B is so definable, which it will be if the set of things we can prove is axiomatic, then the set of things that hold in this language *is* inconsistent. And there are reasons for supposing that this is indeed the case.[152] Even this does not necessarily mean that the familiar natural numbers behave strangely, however. As the model with the "point at infinity" of 9.2 showed, it is quite possible for inconsistent models to have the ordinary natural numbers as a substructure.[153] There are just more possibilities in Heaven and Earth than are dreamt of in a consistent philosophy.

10 PHILOSOPHICAL REMARKS

In previous parts I have touched occasionally on the philosophical aspects of paraconsistency. In this section I want to take up a few of the philosophical implications of paraconsistency at slightly greater length. Its major

[151] The paradox is structurally the same as a paradox often called the 'Knower paradox'. In this, B is interpreted as 'It is known that'. For references and discussion of this paradox and others of its kind, see Priest [1991b].

[152] See Priest [1987], ch. 3. This chapter discusses the connection between Gödel's theorem, the paradoxes of self-reference and dialetheism at greater length.

[153] Though whether the theory of that particular model is axiomatisable is currently unknown.

implication is very simple. As I noted in 3.1, the absolute unacceptability of inconsistency has been deeply entrenched in Western philosophy. It is an assumption that has hardly been questioned since Aristotle. Whilst the law of non-contradiction is a traditional statement of this fact, it is *ECQ* which expresses the real *horror contradictionis*: contradictions explode into triviality. Paraconsistency challenges exactly this, and so questions any philosophical claim based on this supposed unacceptability. This does not mean that consistency cannot play a regulative function: it may still be an expected norm, departure from which requires a justification; but it can no longer provide a constraint of absolute nature. Given the centrality of consistency to Western thought, the philosophical ramifications of paraconsistency are bound to be profound, and this is hardly the place to take them all—or even some—up at great length. What I will do here is consider various objections to employing a paraconsistent logic, and explore a little some of the philosophical issues that arise in this context. In the process we will need to consider not only the purposes of logic, but also the natures of negation, denial, rational belief and belief revision.[154]

10.1 Instrumentalism and Information

Why, then, might one object to paraconsistent logic? Logic has many uses, and any objection to the use of a paraconsistent logic must depend on what it is supposedly being used for. One thing one may want a logic for is to draw out consequences of some information in a purely instrumental way. In such circumstances one may use any logic one likes provided that it gives appropriate results. And if the information is inconsistent, an explosive logic is hardly likely to do this.

Referring back to the list of motivations for the use of a paraconsistent logic in 2.2, drawing inferences from a scientific theory would fall into this category if one is a scientific instrumentalist. Drawing inferences from the information in a computer data base could also fall into this category. If the logic gives the right results—or at least, does not give the wrong results—use it.

The only objection that there is likely to be to the use of a paraconsistent logic in this context is that it is too weak to be of any serious use. One might note, for example, that most paraconsistent logics invalidate the disjunctive syllogism, a special case of resolution, on the basis of which many theorem-provers work.[155] This objection carries little weight, however. Theorem-

[154] Other philosophical aspects of paraconsistency are discussed in numerous places, e.g., da Costa [1982], Priest [1987], Priest *et al.* [1989], ch. 18.

[155] It is worth noting, however, that some theorem-provers that use resolution are not complete with respect to classical semantics. For example, to determine whether α follows from the information in a data base, some theorem-provers employ a heuristic that requires them resolve $\neg\alpha$ with something on the data base, and so on recursively. Em-

provers can certainly be based on other mechanisms.[156] Moreover, the inferential moves of the standard programming language PROLOG can all be interpreted validly in many paraconsistent logics (when ':-' is interpreted as '$\dashv$').

One will often require a logic for something other than merely instrumental use. This does not mean that one is necessarily interested in truth-preservation, however. One might, for example, require a logic whose valid inferences preserve, not truth, but information. The computer case could also be an example of this. Other natural examples of this in the list of 2.2 are the fictional and counterfactual situations. By definition, truth is not at issue here.[157]

Information-preservation implies truth preservation, presumably, but the converse is not at all obvious, and not even terribly plausible. The information that the next flight to Sydney leaves at 3.45 and does not leave at 3.45 would hardly seem to contain the information that there is life on Mars. A paraconsistent logic is therefore a plausible one in this context.

What information, and so information-preservation, are, is an issue that is currently much discussed. One popular approach is based on the situation semantics of Barwise and Perry [1983].[158] This takes a unit of information (an *infon*) to be something of the form $\langle R, a_1, ..., a_n, s\rangle$, where R is an n-place relation, the a_i's are objects, and s is a sign-bit (0 or 1). A situation is a set of infons. The situations in question do not have to be veridical in any sense. In particular, they may be both inconsistent and incomplete. In fact, it is easy to see that a situation, so characterised, is just a relational FDE evaluation. This approach to information therefore naturally incorporates a paraconsistent logic, which may be thought of as a logic of information preservation.[159]

10.2 Negation

Another major use of logic (perhaps the one that many think of first) is in contexts where we want inference to be truth-preserving; for example,

---

ploying this procedure when the data base is $\{p, \neg p\}$ and the query is q will result in a negative answer. Such inference engines are therefore paraconsistent, though they do not answer to any principled semantics that I am aware of.

[156]For details of some automated paraconsistent logics, see, e.g., Blair and Subrahmanian [1988], Thistlewaite *et al.* [1988].

[157]One might also take the other example on that list, constitutions and other legal documents, to be an example of this. Such documents certainly contain information. And one might doubt that this information is the sort of thing that is true or false: it can, after all, be brought into effect by fiat—and may be inconsistent. However, if it is that sort of thing, legal reasoning concerning it would seem to require truth-preservation.

[158]See, e.g., Devlin [1991].

[159]It is worth noting that North American relevant logicians have very often—if not usually—thought of the FDE valuations information-theoretically, as *told* true and *told* false. See, e.g., Anderson *et al.* [1992], sect. 81.

where we are investigating the veridicality of some theory or other. And here, it is very natural to object to the use of a paraconsistent logic. Since truth is never inconsistent a paraconsistent logic is not appropriate.

A paraconsistent logician who thinks that truth is consistent may agree with this, in a sense. We have already seen in 7.6 how a paraconsistent logic, applied to a consistent situation, may give classical reasoning. However, a dialetheist *will* object; not to the need for truth preservation, but to the claim that truth is consistent: some contradictions are true: dialetheias.

This is likely to provoke the fiercest objections. Let me start by dividing these into two kinds: local and global. Global objections attack the possibility of dialetheias on completely general grounds. Local objections, by contrast, attack the claim that some particular claims are dialetheic on grounds specific to the situation concerned.

Let us take the global objections first. Why might one think that dialetheias can be ruled out quite generally, independently of the considerations of any particular case? A first argument is to the effect that a contradiction cannot be true, since contradictions entail everything, and not everything is true. It is clear that in the context where the use of a paraconsistent logic is being defended, this simply begs the question.

Of more substance is the following objection. The truth of contradictions is ruled out by the (classical) account of negation, which is manifestly correct. The amount of substance is only slightly greater here, though: the claim that the classical account of negation is manifestly correct is just plain false.

An account of negation is a *theory* concerning the behaviour of something or other. It is sometimes suggested that it is an account of how the particle 'not', and similar particles in other languages, behaves. This is somewhat naive. Inserting a 'not' does not necessarily negate a sentence. (The negation of 'All logicians do believe the classical account of negation' is not 'All logicians do not believe the classical account of negation'.) And 'not' may function in ways that have nothing to do with negation at all. Consider, e.g.: 'I'm not a Pom; I'm English', where it is connotations of what is said that are being rejected, not the literal truth.

It seems to me that the most satisfactory understanding of an account of negation is to regard it as a theory of the relationship between statements that are contradictories. Note that this by no means rules out a paraconsistent account of negation.[160] Even supposing that we characterise contradictories as pairs of formulas such that at least one must be true and at most one can be true—with which an intuitionist would certainly disagree—it is quite possible to have both $\Box(\alpha \vee \neg\alpha)$ and $\neg\Diamond(\alpha \wedge \neg\alpha)$ valid in a paraconsistent logic, as we saw in 7.3.

Anyway, whatever we take a theory of negation to be a theory *of*, it is but

[160]As Slater [1995] claims.

a *theory*. And different theories are possible. As we have already observed, Aristotle gave an account of this relationship that was quite different from the classical account as it developed after Boole and Frege. And modern intuitionists, too, give a quite different account. Which account is correct is to be determined by the usual criteria for the rational acceptability of theories. (I will say a little more about this later.) The matter is not at all obvious.

Quine is well known for his objection to non-classical logic in general, and paraconsistent logic in particular, on the ground that changing the logic (from the classical one) is 'changing the subject', i.e., succeeds only in giving an account of something else ([1970], p. 81). This just confuses logic, *qua* theory, with logic, *qua* object of theory. Changing one's theory of logic no more changes what it is one is theorising about—in this case, relationships grounding valid reasoning—than changing one's theoretical geometry changes the geometry of the cosmos. Nor does it help to suppose that logic, unlike geometry, is analytic (i.e., true solely in virtue of meanings). Whether or not, e.g., 'There will be a sea battle tomorrow or there will not' is analytic in this sense, is no more obvious than is the geometry of the cosmos. And changing from one theory, according to which it is analytic, to another, according to which it is not, does not change the facts of meaning.[161]

How plausible a paraconsistent account of negation is depends, of course, on which paraconsistent account of negation is given. As we saw in part 4, there are many. One of the simplest and most natural is provided by the relational semantics of 4.5. This is just the classical account, except that classical logic makes the *extra* assumption that all statements have *exactly* one truth value. And logicians as far back as Aristotle have questioned that assumption.[162]

10.3 Denial

Another global objection to dialetheism goes by way of a supposed connection between negation and denial. It is important to be clear about the distinction between these two things for a start. Negation is a syntactic and/or semantic feature of language. Denial is a feature of language use: it is one particular kind of force that an utterance can have, one kind of illocutionary act, as Austin put it. Specifically, it is to be contrasted with assertion.[163] Typically, to assert something is to express one's belief in, or acceptance of, it (or some Gricean sophistication thereof). Typically, to deny something is to express ones rejection of it, that is, one's refusal to ac-

[161]The analogy between logic and geomety is discussed further in Priest [1997a].

[162]The topics of this section and the next are discussed at greater length in Priest [1999].

[163]Traditional logic usually drew the distinction, not in terms of saying, but in terms of judging. It can be found in these terms, for example, in the *Port-Royal Logic* of Arnauld and Nicole.

cept it (or some Gricean sophistication thereof). Clearly, if one is uncertain about a claim, one may wish neither to assert nor to deny it.

Although assertion and denial are clearly different kinds of speech act, Frege argued, and many now suppose, that denial may be reduced to the assertion of negation.[164] If this is correct, then dialetheism faces an obvious problem. Even if some contradictions were true, no one could ever endorse a contradiction, since they could not express an acceptance of one of the contradictories without expressing a rejection of the other.[165]

Frege's reduction has no appeal if we take the negation of a statement simply to be its contradictory. In asserting 'Some men are not mortal', I am not denying 'All men are mortal'. I might not even realise that these are contradictories, and neither might anyone else. And if this does not seem plausible in this simple case, just make the example more complex, and recall that there is no decision procedure for contradictories.

The reduction takes on more plausibility if we identify the negation of a sentence as that sentence prefixed by 'It is not the case that'. But even in this case, the claim that to assert a negation is invariably to deny the sentence negated appears to be false. Dialetheists who asserts both, e.g., 'The Liar sentence is true' and 'It is not the case that the Liar sentence is true', are not expressing the rejection of the former with the latter: they are simply expressing their acceptance of a certain negated sentence.[166] It may well be retorted that this reply just begs the question, since what is at issue is whether a dialetheist *can* do just this. This may be so; but it may now be fairly pointed out that the original objection just begs the question against the dialetheist too.

In any case, there would appear to be plenty of other examples where to assert a negation is not to deny. For example, we may be brought to see that our views are inconsistent by being questioned in Socratic fashion and thus made to assert an explicit contradiction. When this happens, we are not expressing the rejection of any view. What the questioning exposes is exactly our *acceptance* of contradictory views. We may, in the light of the questioning, come to reject one of the contradictories, and so revise our views; but that is another matter.[167]

To assert a negated sentence is not, then, *ipso facto* to deny the sentence negated. Some, having taken this point to heart, object on the other side of the street: dialetheists have no way of expressing some of their views, specifically their rejection of certain claims: we need take nothing a dialetheist

[164] See Frege [1919].

[165] For an objection along these lines, see Smiley in Priest and Smiley [1993].

[166] And even those who take negation to express denial must hold that there is more to the meaning of negation than this. It cannot, for example, perform that function when it occurs attached to *part* of a sentence.

[167] Some non-dialetheists have even argued that it may not even be rational to revise our views in some contexts. See, e.g., Prior on the paradox of the preface [1971], pp. 84f.

says as a denial.[168]

This objection is equally flawed. For a start, even if to assert a negated sentence is to deny it, it is certainly not the only way to deny it. One can do so by a certain shake of the head, or by the use of some other body language. A dialetheist may deny in this way. Moreover, just because the assertion of a negated sentence by a dialetheist (or even a non-dialetheist, as we have seen) may not be a denial, it does not follow that it is not. In denial, a person aims to communicate to a listener a certain mental state, that of rejection; and asserting a negated sentence with the right intonation, and in the right context, may well do exactly that—even if the person is a dialetheist.[169]

10.4 The Rational Acceptability of Contradictions

This does not exhaust the possible global objections to dialetheism,[170] but let us move on to the local ones. These do not object to the possibility of dialetheism in general, but to particular (supposed) cases of it. We noted in 2.2 that a number of these have been proposed, which include legal dialetheias, descriptions of states of change, borderline cases of vague predications and the paradoxes of self-reference. Though the detailed reasons for endorsing dialetheism in each case are different, their general form is the same: a dialetheic account of the phenomenon in question provides the most satisfactory way of handling the problems it poses. A local objection may therefore be provided by producing a consistent account of the phenomenon, and arguing this is rationally preferable. The precise issues involved here will, again, depend on the topic in question; but let us examine one issue in more detail. This will allow the illustration of a number of more general points.

The case we will look at is that of the semantic paradoxes. The background to this needs no long explanation, since a logician or philosopher who does not know it may fairly be asked where they have been this century. Certain arguments such as the Liar paradox, and many others discovered in the middle ages and around the turn of this century, appear to be sound arguments to the effect that certain contradictions employing self-reference and semantic notions are true. A dialetheic account simply endorses the semantic principles in question, and thus the contradictions to which these give rise. A consistent account must find some way of rejecting the reasoning, notably by giving a different account of how the semantic apparatus

[168]Objections along these lines can be found in Batens [1990], and Parsons [1990]. A reply can be found in Priest [1995].

[169]That the same sentence may have different forces in different contexts is hardly a novel observation. For example, an utterance of 'Is the door closed', can be a question, a request or a command, depending on context, intonation, power-relationships, etc.

[170]Some others, together with appropriate discussion, can be found in Sainsbury [1995], ch. 6.

functions. This account must both do justice to the data, and avoid the contradictions.

Many such accounts have, of course, been offered. But they are all well known to suffer from various serious problems. For example, they may provide no independent justification for the restrictions on the semantic principles involved, and so fail to explain why we should be so drawn to the general and contradiction-producing principles. They are often manifestly contrived and/or fly in the face of other well established views. Perhaps most seriously, none of them seems to avoid the paradoxes: all seem to be subject to extended paradoxes of one variety or another.[171] If the global objections to dialetheism have no force, then, the dialetheic position here seems manifestly superior.[172]

It might be said that the inconsistency of the theory is at least a *prima facie* black mark against it. This may indeed be so; but even if one of the consistent theories could find plausible replies to it problems, as long as the theory is complex and fighting a rearguard action, the dialetheic account may still have a simplicity, boldness and mathematical elegance that makes it preferable.

As orthodox philosophy of science realised a long time ago, there are many criteria which are good-making for theories: simplicity, adequacy to the data, preservation of established problem-solutions, etc.; and many which are bad-making: being contrived, handling the data in an *ad hoc* way, and, let us grant, being inconsistent, amongst others.[173] These criteria are usually orthogonal, and may even pull in opposite directions. But when applied to rival theories, the combined effect may well be to render an inconsistent theory rationally preferable to its consistent rival.

General conclusion: a theory in some area is to be rationally preferred to its rivals if it best satisfies the standard criteria of theory choice, familiar from the philosophy of science. An inconsistent theory may be the only viable theory; and even if it is not, it may still, on the whole, be rationally preferable.[174]

---

[171]All this is documented in Priest [1987], ch. 2.

[172]One strategy that may be employed at this point is to argue that a dialetheic theory is trivial, and hence that any other theory, even one with problems, is better. As we have seen, dialetheic truth-theory is non-trivial, but one might nonetheless hope to prove that it is trivial when conjoined with other unobjectionable apparatus. Such arguments have been put forward by Denyer [1989], Smiley, in Priest and Smiley [1993], and Everett [1995] and elsewhere. Replies can be found in, respectively, Priest [1989b], Priest and Smiley [1993], and Priest [1996]. Since my aim here is to illustrate general features of the situation, I will not discuss these arguments.

[173]Though one might well challenge the last of these as a universal rule. There might be nothing wrong with *some* contradictions at all. See Priest [1987], sect. 13.6, and Sylvan [1992], sect. 2.

[174]For a longer discussion of the relationship between paraconsistency and rationality, see Priest [1987], ch. 7.

10.5 Boolean Negation

Another sort of local objection to some dialetheic theories is based on the claim that, whatever one says about negation, there is certainly an operator that behaves in the way that Boolean negation does—call it what you like. Some paraconsistent logicians may even agree with this. (As we saw in 5.3, such an operator is definable in some of the da Costa systems.) And if the point is correct, it suffices to dispose of any dialetheic account of the semantic paradoxes which endorses the T-schema; similarly, any account of set theory that endorses the Comprehension schema. For as I observed in the introduction to part 8, these schemas will then generate Boolean contradictions, and so entail triviality.

Someone who endorses such an account of semantics or set theory must therefore object to the claim that there is an operator that behaves as does Boolean negation. Why, after all, should we suppose this?[175] It might be suggested that we can simply define an operator, $-$, satisfying the proof theoretic principles of Boolean negation, and in particular: $\alpha, -\alpha \vdash \beta$. Such a suggestion would fail: the reason is simply that there is no guarantee that a connective, so characterised, has any determinate sense. The point was made by Prior [1960], who illustrated it with the operator "tonk", $*$, supposedly characterised by the rules $\alpha \vdash \alpha * \beta$, $\alpha * \beta \vdash \beta$. Such an operator induces triviality and can make no sense. Similarly, a paraconsistent logician who endorses the T-schema may fairly point out that the supposition that there is an operator satisfying the proof-theoretic conditions of Boolean negation induces triviality, and so makes no sense.[176]

The claim is theory laden, in the sense that it presupposes that the T-schema is correct. (The addition of such an operator need not produce triviality if only more limited machinery is present.) But any claim about what makes sense is bound to be theory-laden in a similar way. Prior's argument, for example, presupposes the transitivity of deducibility, which may be questioned, as we have seen. The thought that Boolean negation is meaningless may initially be somewhat shocking. But the point has been argued by intuitionist logicians for many years. And though the grounds are quite different,[177] the paraconsistent logician sides with the intuitionist against the classical logician on this occasion.

Can we not, though, characterise Boolean negation semantically, and so show that it is a meaningful connective? The answer is, again, no; not without begging the question. How one attempts to characterise Boolean negation semantically will depend, of course, on one's preferred sort of se-

[175] The following material is covered in more detail in Priest [1990].

[176] There may, of course, be operators that behave like Boolean negation in a limited domain. That is another matter.

[177] The intuitionist reason is that meaningful logical operators cannot generate statements with recognition-transcendent truth conditions, which Boolean negation does. See, e.g., Dummett [1975].

mantics. Let me illustrate the matter with the Dunn semantics. Similar considerations apply to others. With these semantics, the natural attempt to characterise Boolean negation is:

$-\alpha\rho 1$ iff it is not the case that $\alpha\rho 1$
$-\alpha\rho 0$ iff $\alpha\rho 1$

And such a characterisation makes perfectly good semantic sense. However, it does not entail that $-$ satisfies the Boolean proof-theoretic principles. Why should one suppose, crucially, that it validates $\alpha, -\alpha \models \beta$? From the characterisation, it certainly follows that for all ρ, it is not the case that $\alpha\rho 1$ and $-\alpha\rho 1$; but to infer from this that for all ρ, if $\alpha\rho 1$ and $-\alpha\rho 1$ then $\beta\rho 1$ (which states that the inference is valid), just employs the principle of inference that a conditional is true if the negation of its antecedent is. And no sensible paraconsistent conditional validates this.

In other words, to insist that $-$, so characterised, is explosive, just begs the question against the paraconsistentist. And if it is claimed that the negation in the statement of the truth conditions is itself Boolean, and so the inference is acceptable, this again begs the question: whether there is a connective satisfying the Boolean proof-theoretic conditions is exactly what is at issue.[178]

10.6 Logic as an Organon of Criticism

We have now noted three reasons why one might employ a logic: as a purely instrumental means of generating consequences, as an organon of information preservation, and as an organon of truth preservation. This does not exhaust the uses for which one might employ a logic. Another very traditional one is as an organon of criticism, to force others to revise their views. One may object to the use of a paraconsistent logic in this context as follows. If one subscribes to a paraconsistent logic, then there is nothing to stop a person from accepting any inconsistency to which their views lead. Hence, paraconsistency renders logic useless in this context.[179]

The move from the premise that contradictions do not entail everything to the claim that there is nothing to stop a person subscribing to a contradiction is a blatant *non-sequitur*. The threat of triviality may be a reason

[178] I have sometimes heard it said that the logic of a metatheory must be classical. This is just false, as the existence of intuitionist metatheories serves to remind. For certain purposes a dialetheist may, in any case, use a classical metatheory. If, for example, we are trying to show a certain theory to be non-trivial, it suffices to show all the theorems have some property which not all sentences have. This might well be shown using ZF. As we saw in 8.6, ZF makes perfectly good dialetheic sense.

[179] An objection of this kind is to be found in Popper [1963], pp. 316-7. The following is discussed at greater length in Priest [1987], ch. 7.

for revision; it is not the only reason. This is quite obvious in the case of non-dialetheic paraconsistency. If a contradiction is entailed by one's views, then even though they do not explode into triviality, they are still not true. One will still, therefore, wish to revise. One may not, as in the classical case, have to revise immediately. It may not be at all clear how to revise; and in the meantime, an inconsistent but non-trivial belief set is better than no belief set at all. But the pressure will still be there to revise in due course.

The situation may be thought to change if one brings dialetheism into the picture. For the contradiction may then be true, and the pressure to revise is removed. Again, however, the conclusion is too swift. It is certainly true that showing that a person's views are inconsistent may not necessarily force a dialetheist to revise, but other things may well do so. For example, if a person is committed to something of the form $\alpha \rightarrow \perp$, and their views are shown to entail α, there will be pressure to revise, for exactly the classical reason.[180]

Even if a dialetheist's views do not collapse into triviality, the inference to the claim that there is no pressure to revise is still too fast. The fact that there is no *logical* objection to holding on to a contradiction does not show there are no other kinds of objection. There is a lot more to rationality than consistency. Even those who hold consistency to be a constraint on rationality hold that there are many other such constraints. In fact, consistency is a rather weak constraint. That the earth is flat, that Elvis is alive and living in Melbourne, or, indeed, that one is Kermit the Frog, are all views that can be held consistently if one is prepared to make the right kinds of move elsewhere; but these views are manifestly irrational. For a start, there is no evidence for them; moreover, to make things work elsewhere one has to make all kinds of *ad hoc* adjustments to other well-supported views. And whatever constraints there are on rational belief—other than consistency—these work just as much on a dialetheist, and may provide pressure to revise. Not, perhaps, pressure of the ***stand 'em up - knock 'em down*** kind. But such would appear to be illusory in any case. As the history of ideas has shown, rational debates may be a long and drawn out business. There is no magic strategy that will always win the debate—other than employing (or at least showing) the instruments of torture.[181]

[180]Provided that one is not a person who believes that everything is true, then asserting $\alpha \rightarrow \perp$ is a way of denying α. A dialetheist might do this for a whole class of sentences, and so rule out contradictions occurring in certain areas, wholesale.

[181]Avicenna, apparently, realised this. According to Scotus, he wrote that those who deny the law of non-contradiction 'should be flogged or burned until they admit that it is not the same thing to be burned and not burned, or whipped and not whipped'. (*The Oxford Commentary on the Four Books of the Sentences*, Bk. I, Dist. 39. Thanks to Vann McGee for the reference.)

11 CONCLUSION

Let me conclude this essay by trying to put a little perspective into the development of paraconsistent logic. Paraconsistent and explosive accounts of validity are both to be found in the history of logic. The revolution in logic that started around the turn of the century, and which was constituted by the development and application of novel and powerful mathematical techniques, entrenched explosion on the philosophical scene. The application of the same techniques to give paraconsistent logics had to wait until after the second world war.

The period from then until about the late 1970s saw the development of many paraconsistent logics, their proof theories and semantics, and an initial exploration of their possible applications. Though there are still many open problems in these areas, as I have indicated in this essay, the subject was well enough developed by that time to permit the beginning of a second phase: the investigation of inconsistent mathematical theories and structures in their own rights. Whereas the first period was dominated by a negative metaphor of paraconsistency as damage control, the second has been dominated by a more positive attitude: let us investigate inconsistent mathematical structures, both for their intrinsic interest and to see what problems—philosophical, mathematical, or even empirical—they can be used to solve.[182]

Where this stage will lead is as yet anyone's guess. But let me speculate. Traditional wisdom has it that there have been three foundational crises in the history of mathematics. The first arose around the Fourth Century BC, with the discovery of irrational numbers, such as $\sqrt{2}$. It resulted in the overthrow of the Pythagorean doctrine that mathematical truths are exhausted by the domain of the whole numbers (and the rational numbers, which are reducible to these); and eventually, in the development of an appropriate mathematics. The second started in the Seventeenth Century with the discovery of the infinitesimal calculus. The appropriate mathematics came a little faster this time; and the result was the overthrow of the Aristotelian doctrine that truth is exhausted by the domain of the finite (or at least the potential infinite, which is a species of the finite). The third crisis started around the turn of this century, with the discovery of apparently inconsistent entities (such as the Russell set and the Liar sentence) in the foundations of logic and set theory—or at least, with the realisation that such entities could not be regarded as mere curiosities. This provided a major—perhaps the major—impetus for the development of paraconsistent logic and mathematics (as far as it has got). And the philosophical result may be the overthrow of another Aristotelian doctrine: that truth is

[182]It must be said that both stages have been pursued in the face of an attitude sometimes bordering on hostility from certain sections of the establishment logico-philosophical, though things are slowly changing.

exhausted by the domain of the consistent.[183]

University of Queensland, Australia.

BIBLIOGRAPHY

[Anderson and Belnap, 1975] A. Anderson and N. Belnap. *Entailment: the Logic of Relevance and Necessity*, Vol. I, Princeton University Press, Princeton, 1975.

[Anderson *et al.*, 1992] A. Anderson, N. Belnap and J. M. Dunn. *Entailment: the Logic of Relevance and Necessity*, Vol. II, Princeton University Press, Princeton, 1992.

[Avron, 1990] A. Avron. Relevance and Paraconsistency—a New Approach. *Journal of Symbolic Logic*, **55**, 707–732, 1990.

[Arruda, 1977] A. Arruda. On the Imaginary Logic of N. A.Vasil'év', In [Arruda *et al.*, 1977, pp. 3–24].

[Arruda, 1980] A. Arruda. The Paradox of Russell in the System NF_n. In [Arruda *et al.*, 1980b, pp. 1–13].

[Arruda and Batens, 1982] A. Arruda and D. Batens. Russell's Set versus the Universal Set in Paraconsistent Set Theory. *Logique et Analyse*, **25**, 121–133, 1982.

[Arruda *et al.*, 1977] A. Arruda, N. da Costa and R. Chuaqui, eds. *Non-classical Logic, Model Theory and Computability*, North Holland, Amsterdam, 1977.

[Arruda *et al.*, 1980a] A. Arruda, R. Chuaqui and N. da Costa, eds. *Mathematical Logic in Latin America*, North Holland, Amsterdam, 1980.

[Arruda *et al.*, 1980b] A. Arruda, N. da Costa and A. Sette. *Proceedings of the Third Brazilian Conference on Mathematical Logic*, Sociedade Brasiliera de Lógica, São Paulo, 1980.

[Asenjo, 1966] F. G. Asenjo. A Calculus of Antimonies. *Notre Dame Journal of Formal Logic*, **16**, 103–105, 1966.

[Asenjo and Tamburino, 1975] F. G. Asenjo and J. Tamburino. Logic of Antimonies. *Notre Dame Journal of Formal Logic*, **7**, 272–278, 1975.

[Baaz, 1986] M. Baaz. A Kripke-type Semantics for da Costa's Paraconsistent Logic, C_ω. *Notre Dame Journal of Formal Logic*, **27**, 523–527, 1986.

[Barwise and Perry, 1983] J. Barwise and J. Perry. *Situations and Attitudes*, Bradford Books, MIT Press, Cambridge, MA, 1983.

[Batens, 1980] D. Batens. Paraconsistent Extensional Propositional Logics. *Logique et Analyse*, **23**, 195–234, 1980.

[Batens, 1989] D. Batens. Dynamic Dialectical Logics. In [Priest *et al.*, 1989, Chapter 6].

[Batens, 1990] D. Batens. Against Global Paraconsistency. *Studies in Soviet Thought*, **39**, 209–229, 1990.

[Beziau, 1990] J.-Y. Beziau. Logiques Construites Suivant les Méthodes de da Costa, I: Logiques Paraconsistantes, Paracomplètes, Non-Aletheic, Construites Suivant la Premiere Méthode de da Costa. *Logique et Analyse*, **33**, 259–272, 1990.

[Blair and Subrahmanian, 1988] H. A. Blair and V. S. Subrahmanian. Paraconsistent Foundations of Logic Programming. *Journal of Non-Classical Logic*, **5**, 45–73, 1988.

[Błaszczuk, 1984] J. Błaszczuk. Some Paraconsistent Sentential Calculi. *Studia Logica*, **43**, 51–61, 1984.

[Boolos and Jeffrey, 1974] G. Boolos and R. Jeffrey. *Computability and Logic*, Cambridge University Press, Cambridge, 1974.

[183] I am very grateful to those who read the first draft of this essay, and gave me valuable comments and suggestions: Diderik Batens, Ross Brady, Otávio Bueno, Newton da Costa, Chris Mortensen, Nicholas Rescher, Richard Sylvan, Koji Tanaka, Neil Tennant and Matthew Wilson.

[Brady, 1982] R. Brady. Completeness Proofs for the Systems $RM3$ and $BN4$. *Logique et Analyse*, **25**, 9–32, 1982.

[Brady, 1983] R. Brady. The Simple Consistency of Set Theory Based on the Logic CSQ. *Notre Dame Journal of Formal Logic*, **24**, 431–439, 1983.

[Brady, 1989] R. Brady. The Non-Triviality of Dialectical Set Theory. In [Priest *et al.*, 1989, Chapter 16].

[Brady, 1991] R. Brady. Gentzenization and Decidability of Some Contraction-Less Relevant Logics. *Journal of Philosophical Logic*, **20**, 97–117, 1991.

[Brady and Routley, 1989] R. Brady and R. Routley. The Non-Triviality of Extensional Dialectical Set Theory. In [Priest *et al.*, 1989, Chapter 15].

[Brink, 1988] C. Brink. Multisets and the Algebra of Relevant Logics. *Journal of Non-Clasical Logic*, **5**, 75–95, 1988.

[Brown, 1993] B. Brown. Old Quantum Theory: a Paraconsistent Approach. *Proceedings of the Philosophy of Science Association*, **2**, 397–441, 1993.

[Bunder, 1984] M. Bunder. Some Definitions of Negation Leading to Paraconsistent Logics. *Studia Logica*, **43**, 75–78, 1984.

[Bunder, 1986] M. Bunder. Tautologies that, with an Unrestricted Comprehension Schema, Lead to Triviality. *Journal of Non-Classical Logic*, **3**, 5–12, 1986.

[Carnap, 1950] R. Carnap. *Logical Foundations of Probability.* University of Chicago Press, Chicago, 1950.

[Chang, 1963] C. Chang. Logic with Positive and Negative Truth Values. *Acta Philosophica Fennica*, **16**, 19–38, 1963.

[Chellas, 1980] B. Chellas. *Modal Logic: an Introduction*, Cambridge University Press, Cambridge, 1980.

[Curry, 1942] H. Curry. The Inconsistency of Certain Formal Logics. *Journal of Symbolic Logic*, **7**, 115–117, 1942.

[Da Costa, 1974] N. Da Costa. On the Theory of Inconsistent Formal Systems. *Notre Dame Journal of Formal Logic*, **15**, 497–510, 1974.

[Da Costa, 1982] N. Da Costa. The Philosophical Import of Paraconsistent Logic. *Journal of Non-Classical Logic*, **1**, 1–19, 1982.

[Da Costa, 1986] N. Da Costa. On Paraconsistent Set Theory. *Logique et Analyse*, **29**, 361–371, 1986.

[Da Costa and Alves, 1977] N. Da Costa and E. Alves. A Semantical Analysis of the Calculi C_n. *Notre Dame Journal of Formal Logic*, **18**, 621–630, 1977.

[Da Costa and Guillaume, 1965] N. Da Costa and M. Guillaume. Négations Composées et la Loi de Peirce dans les Systems C_n. *Portugaliae Mathematica*, **24**, 201–209, 1965.

[Da Costa and Marconi, 1989] N. Da Costa and D. Marconi. An Overview of Paraconsistent Logic in the 80s. *Journal of Non-Classical Logic* **5**, 45–73, 1989.

[Denyer, 1989] N. Denyer. Dialetheism and Trivialisation. *Mind*, **98**, 259–268, 1989.

[Devlin, 1991] K. Devlin. *Logic and Information*, Cambridge University Press, Cambridge, 1991.

[D'Ottaviano and da Costa, 1970] I. D'Ottaviano and N. da Costa. Sur un Problème de Jaśkowski. *Comptes Rendus Hebdomadaires de l'Academie des Sciences, Paris*, **270A**, 1349–1353, 1970.

[Dowden, 1984] B. Dowden. Accepting Inconsistencies from the Paradoxes. *Journal of Philosophical Logic*, **13**, 125–130, 1984.

[Dummett, 1963] M. Dummett. The Philosophical Significance of Gödel's Theorem. *Ratio*, **5**, 140–55, 1963. Reprinted as [Dummett, 1978, Chapter 12].

[Dummett, 1975] M. Dummett. The Philosophical Basis of Intuitionist Logic. In H. Rose and J.Shepherdson, eds. *Logic Colloquium '73*, North Holland, Amsterdam, 1975. Reprinted as [Dummett, 1978, Chapter 14].

[Dummett, 1977] M. Dummett. *Elements of Intuitionism*, Oxford University Press, Oxford, 1977.

[Dummett, 1978] M. Dummett. *Truth and Other Enigmas*, Duckworth, London, 1978.

[Dunn, 1976] J. M. Dunn. Intuitive Semantics for First Degree Entailment and 'Coupled Trees'. *Philosophical Studies*, **29**, 149–168, 1976.

[Dunn, 1979] J. M. Dunn. A theorem in 3-valued model theory with connections to number theory, type theory and relevance. *Studia Logica*, **38**, 149–169, 1979.

[Dunn, 1980] J. M. Dunn. A Sieve for Entailments. *Journal of Philosophical Logic*, **9**, 41–57, 1980.

[Dunn, 1988] J. M. Dunn. The Impossibility of Certain Second-Order Non-Classical Logics with Extensionality. In *Philosophical Analysis*, D. F. Austin, ed., pp. 261-79. Kluwer Academic Publishers, Dordrecht, 1988.

[Everett, 1995] A. Everett. Absorbing Dialetheias. *Mind*, **103**, 414–419, 1995.

[Frege, 1919] G. Frege. Negation. *Beiträge zur Philosophie des Deutschen Idealisimus*, **1**, 143–157, 1919. Reprinted in translation in *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black, eds., pp. 117–135. Basil Blackwell, Oxford, 1960.

[Goodman, 1981] N. D. Goodman. The Logic of Contradiction. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **27**, 119–126, 1981.

[Goodship, 1996] L. Goodship. On Dialethism. *Australasian Journal of Philosophy*, **74**, 153–161, 1996.

[Haack, 1974] S. Haack. *Deviant Logic*, Cambridge University Press, Cambridge, 1974.

[Halpern, 1986] J. Y. Halpern, ed. *Theoretical Aspects of Reasoning about Knowledge*, Morgan Kaufmann, Los Altos, 1986.

[Jaśkowski, 1969] S. Jaśkowski. Propositional Calculus for Contradictory Deductive Systems. *Studia Logica*, **24**, 143–157, 1969.

[Kaye, 1991] R. Kaye. *Models of Peano Arithmetic*, Clarendon Press, Oxford, 1991.

[Kotas and da Costa, 1978] J. Kotas and N. da Costa. On the Problem of Jaśkowski and the Logics of Łukasiewicz. In [Arruda *et al.*, 1977, pp. 127–139].

[Kotas and da Costa, 1989] J. Kotas and N. da Costa. Problems of Modal and Discussive Logic. In [Priest *et al.*, 1989, Chapter 8].

[Loparić, 1986] A. Loparić. A Semantical Study of some Propositional Calculi. *Journal of Non-Classical Logic*, **3**, 73–95, 1986.

[Loparić and da Costa, 1984] A. Loparić and N. da Costa. Paraconsistency, Paracompleteness and Valuations. *Logique et Analyse*, **27**, 119–131, 1984.

[Lucas, 1961] J. R. Lucas. Minds, Machines and Gödel. *Philosophy*, **36**, 112–127, 1961. Reprinted in *Minds and Machines*, A. Anderson, ed., pp. 43-59. Prentice Hall, Englewood Cliffs, 1964.

[Łukasiewicz, 1971] J. Łukasiewicz. On the Principle of Contradiction in Aristotle. *Review of Metaphysics*, **24**, 485–509, 1971.

[Marconi, 1984] D. Marconi. Wittgenstein on Contradiction and the Philosophy of Paraconsistent Logic. *History of Philosophy Quarterly*, **1**, 333–352, 1984.

[Martin, 1986] C. Martin. William's Machine. *Journal of Philosophy*, **83**, 564–572, 1986.

[Meyer, 1978] R. K. Meyer. Relevant Arithmetic. *Bulletin of the Section of Logic, Polish Academy of Sciences*, **5**, 133–137, 1978.

[Meyer and Mortensen, 1984] R. K. Meyer and C. Mortensen. Inconsistent Models for Relevant Arithmetics. *Journal of Symbolic Logic*, **49**, 917–929, 1984.

[Meyer and Routley, 1972] R. K. Meyer and R. Routley. Algebraic Analysis of Entailment, I. *Logique et Analyse*, **15**, 407–428, 1972.

[Meyer *et al.*, 1979] R. K. Meyer, R. Routley and J. M. Dunn. Curry's Paradox. *Analysis*, **39**, 124–128, 1979.

[Mortensen, 1980] C. Mortensen. Every Quotient Algebra for C_1 is Trivial. *Notre Dame Journal of Formal Logic*, **21**, 694–700, 1980.

[Mortensen, 1984] C. Mortensen. Aristotle's Thesis in Consistent and Inconsistent Logics. *Studia Logica*, **43**, 107–116, 1984.

[Mortensen, 1987] C. Mortensen. Inconsistent Nonstandard Arithmetic. *Journal of Symbolic Logic*, **52**, 512–518, 1987.

[Mortensen, 1989] C. Mortensen. Anything is Possible. *Erkenntnis*, **30**, 319–337, 1989.

[Mortensen, 1995] C. Mortensen. *Inconsistent Mathematics*, Kluwer Academic Publishers, Dordrecht, 1995.

[Nelson, 1959] D. Nelson. Negation and Separation of Concepts in Constructive Systems. In *Constructivity in Mathematics*, A. Heyting, ed., pp. 208–225 North Holland, Amsterdam, 1959.

[Parsons, 1990] T. Parsons. True Contradictions. *Canadian Journal of Philosophy*, **20**, 335–353, 1990.

[Peña, 1984] L. Peña. Identity, Fuzziness and Noncontradiction. *Noûs*, **18**, 227–259, 1984.

[Peña, 1989] L. Peña. Verum et Ens Convertuntur. In [Priest *et al.*, 1989, Chapter 20].

[Popper, 1963] K. R. Popper. *Conjectures and Refutations*, Routledge and Kegan Paul, London, 1963.

[Prawitz, 1965] D. Prawitz. *Natural Deduction*, Almqvist & Wiksell, Stockholm, 1965.

[Priest, 1979] G. Priest. Logic of Paradox. *Journal of Philosophical Logic*, **8**, 219–241, 1979.

[Priest, 1980] G. Priest. Sense, Truth and *Modus Ponens*. *Journal of Philosophical Logic*, **9**, 415–435, 1980.

[Priest, 1982] G. Priest. To Be and Not to Be: Dialectical Tense Logic. *Studia Logica*, **41**, 249–268, 1982.

[Priest, 1987] G. Priest. *In Contradiction*, Martinus Nijhoff, the Hague, 1987.

[Priest, 1989a] G. Priest. Dialectic and Dialetheic. *Science and Society*, **53**, 388–415, 1989.

[Priest, 1989b] G. Priest. Denyer's $ Not Backed by Sterling Arguments. *Mind*, **98**, 265–268, 1989.

[Priest, 1990] G. Priest. Boolean Negation and All That. *Journal of Philosophical Logic*, **19**, 201–215, 1990.

[Priest, 1991a] G. Priest. Minimally Inconsistent *LP*. *Studia Logica*, **50**, 321–331, 1991.

[Priest, 1991b] G. Priest. Intensional Paradoxes. *Notre Dame Journal of Formal Logic*, **32**, 193–211, 1991.

[Priest, 1992] G. Priest. What is a Non-Normal World? *Logique et Analyse*, **35**, 291–302, 1992.

[Priest, 1994] G. Priest. Is Arithmetic Consistent? *Mind*, **103**, 321–331, 1994.

[Priest, 1995] G. Priest. Gaps and Gluts: Reply to Parsons. *Canadian Journal of Philosophy*, **25**, 57–66, 1995.

[Priest, 1996] G. Priest. Everett's Trilogy. *Mind*, **105**, 631–647, 1996.

[Priest, 1997a] G. Priest. On a Paradox of Hilbert and Bernays. *Journal of Philosophical Logic*, **26**, 45–56, 1997.

[Priest, 1997b] G. Priest. Inconsistent Models of Arithmetic: Part I, Finite Models. *Journal of Philosophical Logic*, **26**, 223–235, 1997.

[Priest, 1998] G. Priest. To be *and* Not to Be — That is the Answer. On Aristitle on the Law of Non-contradiction. *Philosophiegeschichte und Logische analyse*, **1**, 91–130, 1998.

[Priest, 1999] G. Priest. What Not? A Defence of a Dialetheic Theory of Negation. In *Negation*, D. Gabbay and H. Wansing, eds., pp. 101–120. Kluwer Academic Publishers, Dordrecht, 1999.

[Priest, forthcoming:a] G. Priest. On Alternative Geometries, Arithmetics and Logics; a Tribute to Łukasiewicz', In *Proceedings of the Conference Łukasiewicz in Dublin*, M. Baghramian, ed., to appear.

[Priest, forthcoming:b] G. Priest. Inconsistent Models of Arithmetic: Part II, the General Case. *Journal of Symbolic Logic*, to appear.

[Priest *et al.*, 1989] G. Priest, R. Routley and G. Norman, eds. *Paraconsistent Logic: Essays on the Inconsistent*, Philosophia Verlag, Munich, 1989.

[Priest and Smiley, 1993] G. Priest and T. Smiley. Can Contradictions be True? *Proceedings of the Aristotelian Society, Supplementary Volume*, **65**, 17–54, 1993.

[Priest and Sylvan, 1992] G. Priest and R. Sylvan. Simplified Semantics for Basic Relevant Logics. *Journal of Philosophical Logic*, **21**, 217–232, 1992.

[Prior, 1960] A. Prior. The Runabout Inference Ticket. *Analysis*, **21**, 38–39, 1960. Reprinted in *Philosophical Logic*, P. Strawson, ed. Oxford University Press, Oxford, 1967.

[Prior, 1971] A. Prior. *Objects of Thought*, Oxford University Press, Oxford, 1971.

[Pynko, 1995a] A. Pynko. Characterising Belnap's Logic via De Morgan Laws. *Mathematical Logic Quarterly*, **41**, 442–454, 1995.

[Pynko, 1995b] A. Pynko. Priest's Logic of Paradox. *Journal of Applied and Non-Classical Logic*, **2**, 219–225, 1995.

[Quine, 1970] W. V. O. Quine. *Philosophy of Logic*, Prentice-Hall, Englewood Cliffs, 1970.

[Rescher, 1964] N. Rescher. *Hypothetical Reasoning*, North Holland Publishing Company, Amsterdam, 1964.

[Rescher, 1969] N. Rescher. *Many-valued Logic*, McGraw-Hill, New York, 1969.

[Rescher and Brandom, 1980] N. Rescher and R. Brandom. *The Logic of Inconsistency*, Basil Blackwell, Oxford, 1980.

[Rescher and Manor, 1970–71] N. Rescher and R. Manor. On Inference from Inconsistent Premises. *Theory and Decision*, **1**, 179–217, 1970–71.

[Restall, 1992] G. Restall. A Note on Naive Set Theory in *LP*. *Notre Dame Journal of Formal Logic*, **33**, 422–432, 1992.

[Restall, 1993] G. Restall. Simplified Semantics for Relevant Logics (and Some of Their Rivals). *Journal of Philosophical Logic*, **22**, 481–511, 1993.

[Restall, 1995] G. Restall. Four-Valued Semantics for Relevant Logics (and Some of Their Rivals). *Journal of Philosophical Logic*, **24**, 139–160, 1995.

[Rosser and Turquette, 1952] J. B. Rosser and A. R. Turquette. *Many-valued Logics*, North Holland, Amsterdam, 1952.

[Routley, 1978] R. Routley. Semantics for Connexive Logics, I. *Studia Logica*, **37**, 393–412, 1978.

[Routley, 1979] R. Routley. Dialectical Logic, Semantics and Metamathematics. *Erkenntnis*, **14**, 301–331, 1979.

[Routley, 1980a] R. Routley. Problems and Solutions in Semantics of Quantified Relevant Logics. In [Arruda *et al.*, 1980a, pp. 305–340].

[Routley, 1980b] R. Routley. Ultralogic as Universal. Appendix I of *Exploring Meinong's Jungle and Beyond*, Research School of Social Sciences, Australian National University, Canberra, 1980.

[Routley, 1984] R. Routley. The American Plan Completed; Alternative Classical-Style Semantics, without Stars, for Relevant and Paraconsistent Logics. *Studia Logica*, **43**, 131–158, 1984.

[Routley, 1989] R. Routley. Philosophical and Linguistic Inroads: Multiply Intensional Relevant Logics. In *Directions in Relevant Logic*, J. Norman and R. Sylvan, eds. Ch. 19. Kluwer Academic Publishers, Dordrecht, 1989.

[Routley and Loparić, 1978] R. Routley and A. Loparić. A Semantical Analysis of Arruda–da Costa *P* Systems and Adjacent Non-Replacement Systems. *Studia Logica*, **37**, 301–320, 1978.

[Routley and Loparić, 1980] R. Routley and A. Loparić. Semantics for Quantified and Relevant Logics without Replacement. In [Arruda *et al.*, 1980b, pp. 263–280].

[Routley and Meyer, 1973] R. Routley and R. K. Meyer. The Semantics of Entailment, I. In *Truth, Syntax and Modality*, H. Leblanc, ed. North Holland, Amsterdam, 1973.

[Routley *et al.*, 1982] R. Routley, V. Plumwood, R. K. Meyer and R. Brady. *Relevant Logics and Their Rivals*, Vol. I, Ridgeview, Atascadero, 1982.

[Routley and Routley, 1972] R. Routley and V. Routley. The Semantics of First Degree Entailment. *Noûs*, **6**, 335–359, 1972.

[Routley and Routley, 1989] R. Routley and V. Routley. Moral Dilemmas and the Logic of Deontic Notions. In [Priest *et al.*, 1989, Chapter 23].

[Sainsbury, 1995] M. Sainsbury. *Paradoxes*, (2nd ed.), Cambridge University Press, Cambridge, 1995.

[Schotch and Jennings, 1980] P. Schotch and R. Jennings. Inference and Necessity. *Journal of Philosophical Logic*, **9**, 327–340, 1980.

[Slaney, 1989] J. Slaney. *RWX* is not Curry Paraconsistent. In [Priest *et al.*, 1989, Chapter 17].

[Slater, 1995] B. Slater. Paraconsistent Logics? *Journal of Philosophical Logic*, **24**, 451–454, 1995.

[Smiley, 1959] T. J. Smiley. Entailment and Deducibility. *Proceedings of the Aristotelian Society*, **59**, 233–254, 1959.

[Sylvan, 1992] R. Sylvan. Grim Tales Retold: How to Maintain Ordinary Discourse about—and Despite—Logically Embarrassing Notions and Totalities. *Logique et Analyse*, **35**, 349–374, 1992.

[Sylvan, 2000] R. Sylvan. A preliminary western history of sociative logics. In *Sociative Lgoics and hteir Applications: Essays by the late Richard Sylvan*, D. Hyde and G. Priest, eds. Ashgate Publishers, Aldershot, 2000.

[Tennant, 1980] N. Tennant. A Proof-Theoretic Approach to Entailment. *Journal of Philosophical Logic*, **9**, 185–209, 1980.

[Tennant, 1984] N. Tennant. Perfect Validity, Entailment and Paraconsistency. *Studia Logica*, **43**, 181–200, 1984.

[Tennant, 1987] N. Tennant. *Anti-Realism and Logic: Truth as Eternal*, Clarendon Press, Oxford, 1987.

[Tennant, 1992] N. Tennant. *Autologic*, Edinburgh University Press, Edinburgh, 1992.

[Thistlewaite *et al.*, 1988] P. Thistlewaite, M. McRobbie and R. K. Meyer. *Automated Theorem-Proving in Non-Classical Logics*, John Wiley & Sons, New York, 1988.

[Urbas, 1989] I. Urbas. Paraconsistency and the C-systems of da Costa. *Notre Dame Journal of Formal Logic*, **30**, 583–597, 1989.

[Urbas, 1990] I. Urbas. Paraconsistency. *Studies in Soviet Thought*, **39**, 343–354, 1990.

[Urquhart, 1984] A. Urquhart. The Undecidability of Entailment and Relevant Implication. *Journal of Symbolic Logic*, **49**, 1059–1073, 1984.

[White, 1979] R. White. The Consistency of the Axiom of Comprehension in the Infinite-Valued Predicate Logic of Łukasiewicz. *Journal of Philosophical Logic*, **8**, 509–534, 1979.

[Wittgenstein, 1975] L. Wittgenstein. *Philosophical Remarks*, Basil Blackwell, Oxford, 1975.