39

On Paraconsistency

BRYSON BROWN

1 What is Paraconsistency?

The term paraconsistent logic is due to F. Miro Quesada. 'Para' can mean any of 'against,' 'near,' or 'beyond.' It's not clear which of these Quesada had specifically in mind, or whether he deliberately embraced the ambiguities inherent in the term. But each of these meanings suits the programs of at least some who have worked on paraconsistent logic. The most radical paraconsistentists, the dialetheists, are indeed against consistency, at least as a global constraint on our metaphysics. They hold that the world is inconsistent, and aim at a general logic that goes beyond all the consistency constraints of classical logic. More modest practitioners of paraconsistent logic aim to give us a logic suitable for treating nearly (but not quite) consistent sets. But the central logical problem is something both the moderates and the radicals share: the classical trivialization of inconsistent and unsatisfiable sets.

A set Γ is inconsistent iff its closure under deduction includes both α and $\neg\alpha$ for some sentence α; it is unsatisfiable if there is no admissible valuation that satisfies all members of Γ. Classical logic trivializes all such premise sets. That is, from any inconsistent premise set, we can derive any sentence in the language, and from any unsatisfiable premise set, every sentence in the language follows. Each term when I teach introductory logic I have to explain this intuitively odd fact, which we can express roughly as:[1]

Triv: $\{A,\neg A\} \vdash_c B$; $\{A,\neg A\} \models_c B$

In that logic class I sometimes try to motivate this oddity by locating it in the context of persuading others to accept various sentences. Of course we know the premise set $\{A,\neg A\}$ cannot be true. So, if someone grants you (or anyone) those premises, they should be prepared to grant you anything at all (how could they object to B, having already accepted A and $\neg$A?). But this defense is just a rhetorical dodge. It assumes something that's clearly false, that is that we can never have good reasons to accept both A and $\neg$A, while not having any good reason to accept B.

More famously, C. I. Lewis argued for Triv by presenting a proof of it, along these lines:

1.	A	premise
2.	$\neg A$	premise
3.	$(A \vee B)$	1, $\vee$-intro
4.	B	2,3, disjunctive syllogism

The proof produces a challenge for those who reject Triv: Which of these two rules, $\vee$-intro or disjunctive syllogism, will they give up? Both seem pretty obviously sound. How can we deny that $A \vee B$ follows from A, given the standard understanding of '$\vee$' as inclusive disjunction? Similarly, how can we deny that B follows from $A \vee B$ and $\neg A$, given that '$\neg$' is negation and '$\vee$' is inclusive disjunction? The usual answer (first proposed by Anderson and Belnap[2]) is that disjunctive syllogism must fail in such cases. After all, we obtained $A \vee B$ by inferring it from our premise A. How can we then justify turning around and inferring B with the help of our other premise, $\neg A$? This inference depends on the assumption that the assertion of $\neg A$ can be used to rule out that it's due to having asserted A that we are entitled to assert that $A \vee B$. But to assume this is to deny that we ever reason with sets like $\{A, \neg A\}$, not to infer something we would want to infer from such sets.

There are other reasons to adopt paraconsistent logic, reasons that are telling even for those who accept the argument for trivialization above. For example, a deontic logic that does not trivialize conflicting obligations, or an epistemic logic that does not trivialize inconsistent beliefs, will require a paraconsistent consequence relation. And many philosophers who accept Lewis' trivialization argument are still prepared to accept nontrivial but inconsistent obligations and/or beliefs. Inconsistent theories, such as Bohr's theory of the hydrogen atom, and the early calculus, provide another application that calls for consequence relations that don't trivialize inconsistent sets of sentences. We'll examine some logics specifically designed for these applications later; for now, we will focus on Triv and the view of consequence relations that it derives from. Our first aim will be a general taxonomy of paraconsistent logic based on alternative approaches to avoiding Triv.

2 Motives for Paraconsistency

Many philosophers follow Lewis, and respond to Triv by accepting and defending it. One of the principal arguments in favour of Triv is to say that Triv is just classical logic's way of telling you to get yourself a new set of premises. If your premises are unsatisfiable or inconsistent, then they cannot all be true, and you cannot accept them all without at least implicitly contradiction yourself. But if you know that you are reasoning from a starting point that already contains an error, you should clear that error up, not compound it by drawing further conclusions from your already erroneous premises.

However, as many have pointed out, eliminating inconsistency from our premises is easier said than done. There is a clear technical sense in which we can say that eliminating inconsistency is difficult. Of course, before we can even try to eliminate an inconsistency from our premises, we must find the inconsistency. But the consistency of a set of sentences is not decidable, that is there is no procedure which will show whether or not an arbitrary set of sentences in a first-order language is consistent.[3] So inconsistencies may lie hidden in our set of accepted claims.

More importantly, there may be strong practical reasons for not eliminating an inconsistency, even when we can see that one is present. For example, consider Bohr's model of the hydrogen atom. While Bohr used classical electrodynamics to model the radiation the atom absorbs and emits, he also allowed the electron to orbit the nucleus in a 'stationary state' without emitting any radiation and without the radiation it emits in shifting between stationary states having a frequency related to the frequency of the electron's periodic motion around the nucleus. But classical electrodynamics requires that all accelerated charges radiate, and that the radiation emitted by an accelerated charge have a frequency related to the frequency of the accelerated motion. Bohr's model was clearly inconsistent – but these inconsistent features were indispensable to capturing the phenomena Bohr was trying to understand. The upshot was an extended period in which our best physical theories were clearly inconsistent. A similar situation exists today, since our best theory of the structure of space–time, Einstein's general theory of relativity, is inconsistent with our quantum-mechanical view of microphysics.

Of course the importance of Triv depends on what we take to be the aim of logic, and in particular that of consequence relations. One influential view is that consequence relations are models of inference. By inference, we mean the process of adding new sentences to those we have already accepted by reasoning (rather than by observation). But it's clear that we would never accept every sentence. So if we ever accept inconsistent premises as a starting point for inference, we must use a paraconsistent logic in our inferences. Of course a defender of Lewis' position might argue that we never really accept inconsistent premises. But this seems untenable. After all, we are finite thinkers who do not always see the consequences of everything we accept. In particular, as we saw above, we may well not know our premises are inconsistent.

However, the view that consequence relations aim to model inference is hard to defend. First, as Gilbert Harman has argued, inference (in the sense of reasoning-driven change of belief) does not always involve adding to our set of accepted sentences. Sometimes, when we discover that some new sentence follows from the sentences we have already accepted, our response is to reject one or more of our premises rather than accept the new sentence. Second, real inference is constrained in ways that logical consequence relations are not. Many inferences that are 'logical' in the sense that they preserve truth and consistency are nevertheless a pointless waste of time for anyone to carry out. For example, disjoining a sentence with itself, and conjoining a sentence to itself, are perfectly correct from a merely logical point of view. But it would be ridiculous to waste our time idly inferring such consequences of the sentences we accept.

This leads to an alternative take on what the role of logic is. Inference is a highly pragmatic process involving both logical considerations and practical constraints of salience, along with rich evaluations of how best to respond both to our observations and to the consequences of what we have already accepted. So the role of logical consequence relations is clearly not to tell us what we may or should infer from the sentences we accept. But consequence relations can still tell us what is involved in accepting those sentences, that is they can operate as closure conditions on our commitments. What I mean by this is that, in accepting a set of sentences, we can fairly be said to be *committed* to the consequences of those sentences – including the consequences we would never infer, whether because inferring them would be a waste of time, or because, rather than infer them, we would reject something that we now accept. In fact, I believe it's because we regard our commitments as closed under logical consequence that we are inclined to either accept sentences we find 'follow' logically from our commitments, or to reject some of the commitments they follow from.

This fits both our limits and our intuitions quite well. On the one hand, it does not assume inference is monotonic, or that a proper logic must shape its consequence relation to fit the exigencies of actual inferential practice. But on the other hand, it allows a substantial role for logical consequences in determining the content of our commitments. Without some such closure operation on commitment, our commitments would be too closely tied to the actual sentences we were willing to utter. And the point of assertion is not just to utter a sentence, but to express a commitment to what that sentence says.

This view of consequence relations provides a conclusive argument for paraconsistency. If we ever need to model commitments to inconsistent (and/or unsatisfiable) sets of sentences, then we will need a paraconsistent logic to do it. Otherwise, such commitments will be trivial, and indistinguishable from each other. For example, consider

1. Leibniz' calculus. This mathematical theory is inconsistent because it treats infinitesimals as both equal to zero and not equal to zero. From an addition equation in which an infinitesimal is an addend we can infer the corresponding equation without the infinitesimal, but division by infinitesimals is still well defined.
2. Naïve set theory. This mathematical theory is inconsistent because it claims that there exists a set corresponding to each condition on membership we can state. As a result, it claims that the Russell set exists, *viz*. the set of all sets that are not members of themselves. But if this set is a member of itself, then by its condition of membership, it is not a member of itself. And if it is not a member of itself, then by its condition of membership, it is a member of itself. So it is a member of itself if and only if it is not a member of itself, that it is an inconsistent object.

If we use classical logic to model commitment to these theories, they are equivalent (assuming we employ the same language to present them), and commitment to either is trivial. Avoiding triviality and distinguishing commitment to Leibniz' calculus from commitment to naïve set theory will require a paraconsistent logic.

In a closely connected use of logic, it is standard to call a set of sentences closed under a consequence relation a *theory*. If we are to cope constructively with inconsistent theories, the consequence relation they are closed under will have to be a paraconsistent one. Further applications for paraconsistent logic arise in epistemic logic, when we aim to model inconsistent beliefs, and in deontic logic, when we aim to model inconsistent obligations or rules. Standard modal techniques in epistemic logic represent a belief state in terms of the sentences true at all worlds at which the subject's beliefs are all true. Similarly, standard modal techniques in deontic logic represent obligations by appeal to the sentences true at all the worlds at which those obligations are met, and the demands of rules in a situation by appeal to the sentences true at all (later) situations at which the rules have been obeyed. The result, if these worlds and situations provide classical valuations of the sentences of our language, is that beliefs, obligations, and the demands of rules are represented by classical theories. Coping with inconsistent beliefs, obligations, and rules will require instead that we represent them by paraconsistent theories, that is as sets of sentences closed under a paraconsistent consequence relation.

A more radical motivation for paraconsistency can be found in the work of the Australian dialetheists. The motives we have considered so far are to provide a coherent account of how to reason from inconsistent premises, or how to describe the content of an inconsistent theory. But the dialetheists aim at a logic that will allow that such theories are *true*. In particular, the dialetheists take paradoxes such as the liar and the paradoxes of naïve set theory at face value. That is, they view these paradoxes as proofs that certain inconsistencies are true. Consider the liar paradox:

L: This sentence is false.

The dialetheists claim that L is both true and false. We can prove this by *reductio* – suppose, for *reductio*, that L is true. Then what it says must be true. But it says that L is false. Therefore, if L is true, L must be false. Now suppose for *reductio* that L is false. But L says that L is false. Thus what L says is true, so L is true. Thus if L is false, then L is true. Therefore, L is true if and only if L is false. But if we accept the (semantic) law of excluded middle, that is that every sentence is either true or false, it follows that L must be both.[4]

Whether we want to reason in a non-trivializing way from inconsistent premises, or to model inconsistent beliefs, obligations, or theories, or to reason about an inconsistent metaphysics in which the liar sentence is both true and not true (and the Russell set both is and is not a member of itself), we will need a paraconsistent logic. Of course the particular kind of paraconsistent logic we will choose may depend on which of these applications we have in mind. But we will address the problem of choosing between various paraconsistent logics below.

3 The Sources of Trivialization

The usual definition of validity says that a sentence α can be validly inferred from a set of sentences Γ when and only when α is a semantic consequence of Γ, that is iff $\Gamma \models \alpha$, where:

(1) $\Gamma \models \alpha$ iff the truth of all Γ's members guarantees the truth of α.

This definition provides us with a straightforward explanation of the importance of deductive validity in evaluating arguments. Ideally, an argument should give conclusive reasons in support of its conclusion. And part of what we ask of conclusive reasons is that if the premises (the 'starting points' of reasoning) all hold, then the conclusion will be guaranteed to hold as well. (Of course 'holding' here is just being true.) But as introductory logic courses point out very early on, it is difficult, if not impossible, to say in general when an argument in a natural language is valid. To give such a general account of validity, we would need a theory that identifies all the various sentences of a natural language and then tells us, in a logically revealing way, what it takes for each sentence to be true. Only then could we give general rules for determining whether meeting the truth conditions for an arbitrary set of sentences will guarantee that the truth condition for some other sentence is also met. Such a theory is far beyond us.

Logic courses that seek, nevertheless, to move towards such a theory quickly turn away from arguments expressed in English, French, Japanese, etc., to arguments expressed using formal languages with formal semantics. Such languages provide a formal syntax telling us what strings of symbols count as sentences, and a formal semantics specifying the truth (or 'satisfaction') conditions for sentences in the language. With these in hand, we can say clearly what it would take for a set's members to be true, and then use formal methods to show whether arranging things so that they are true will guarantee a conclusion sentence is true as well.

But there is another approach to capturing the notion of a valid argument, founded in what we regard as simple rules of good reasoning. On this approach, we are given a set of rules to follow in reasoning with sentences. These rules are based on the syntactic structure of the sentences, that is the symbols and how they are arranged in each sentence, rather than on an account of their truth conditions. A sequence of sentences (sometimes accompanied by other book-keeping devices) that follows the rules is called a *derivation*. The rules determine what sentences we may write down where in a derivation, and when a derivation is complete. The aim of this approach is to provide a set of simple, obviously correct steps that are collectively sufficient to capture all the consequences that follow from any premise set. This approach gives rise to the following account of a consequence relation:

(2) $\Gamma \vdash \alpha$ iff α can be derived from Γ.

One advantage of this approach to consequence relations is that it focuses our attention on the process of reasoning, rather than on 'meanings' that are taken to lie behind that process. Either way, as we will see, such consequence relations suffer from a serious limitation. In both cases, a tacit assumption is made about the premise sets we are thinking of – an assumption whose failure must trivialize these consequence relations.

We say that Γ is consistent if and only if it is impossible to derive a sentence and its negation from Γ,[5] and that Γ is satisfiable if and only if some valuation assigns a designated value to every sentence in Γ. A set is *maximally* consistent if it is consistent and adding any sentence to it would render it inconsistent. Similarly, a set is maximally satisfiable if it is satisfiable and no proper superset is satisfiable. The standard account of consequence relations results from a certain picture of what our consequence relations ($\models$ and $\vdash$) are supposed to preserve. Crudely, $\models$ is said to preserve truth while $\vdash$ preserves consistency. More carefully put, $\models$ preserves the satisfiability of all satisfiable extensions of the premise set, and $\vdash$ preserves the consistency of all consistent extensions, in the following senses:

(3) $\Gamma \models \alpha$ iff $\forall\Gamma'[(\Gamma' \supset \Gamma \;\&\; \Gamma'$ is satisfiable$) \rightarrow \Gamma',\alpha$ is satisfiable].
(4) $\Gamma \vdash \alpha$ iff $\forall\Gamma'[(\Gamma' \supset \Gamma \;\&\; \Gamma'$ is consistent$) \rightarrow \Gamma',\alpha$ is consistent].[6]

That is, if Γ' extends Γ satisfiably or consistently, then every consequence of Γ must satisfiably or consistently extend Γ'. We can say that the consequence relations $\models$ and $\vdash$ *beg no questions*, in the sense that closing Γ under $\models$ or $\vdash$adds nothing to Γ that is incompatible with any semantically or syntactically acceptable extensions of Γ. And of course, given the soundness and completeness of our system of derivation (3) and (4) are simply alternative definitions of the same consequence relation.

The principal point of this section can now be set out. If Γ is unsatisfiable (or inconsistent), then Γ has no satisfiable (consistent) extensions. You can never make a set that is already unsatisfiable or inconsistent into a satisfiable or consistent set by adding sentences to it. Thus every sentence trivially 'preserves' the satisfiability or consistency of all satisfiable or consistent extensions of such sets. So clauses (3) and (4) are satisfied for every sentence α, that is, every sentence is a consequence of an unsatisfiable or inconsistent set. The trivialization of unsatisfiable/inconsistent sets of sentences is deeply embedded in these standard accounts of consequence relations.

A logic is (minimally) paraconsistent iff it resists this trivialization, that is iff for some *classically* unsatisfiable or inconsistent set of sentences, the closure of the set under the logic's consequence relation is not the set of all sentences. But there are different ways to go about producing such a logic, rooted in different choices about what we choose to change in (3) and (4).

4 A Natural Taxonomy for Paraconsistent Logics

For a variety of reasons, including a focus on axiomatic presentations of logical systems and the vivid appeal of citing non-Euclidean geometries as a precedent, Jan Łukasiewicz, Jaskowski, and some other early figures saw paraconsistent logic (and nonclassical logic in general) on analogy with non-Euclidean geometry. Certain 'logical laws,' they suggested, could be treated as analogous to Euclid's parallel postulate in geometry, that is as characteristic of a particular *kind* of logic, but not essential for all logic.[7] This approach suggests a taxonomy of paraconsistent logics based on the classical theorems and rules that they retain, and those they give up.

In accord with this, the best known contemporary taxonomy of paraconsistent logic focuses on the tactics by which various paraconsistent logics avoid Triv, that is, the changes that are made to classical axioms and/or derivation rules. In Priest et al. (1989), three principal groups of paraconsistent logic are distinguished. First, Jaskowski's discursive logics are grouped with Rescher and Brandom's semantics for truth-value gaps and gluts and Schotch and Jennings' weakly aggregative logics, under the heading 'non-adjunctive' logics (Priest et al. 1989, p. 57). These logics block the inference from $\{p, \neg p\}$ to $p \wedge \neg p$. While classical contradictions, such as $p \wedge \neg p$, continue to be trivial for these logics, inconsistent sets that contain no contradictions can have non-trivial consequences. Second, the C-systems of Da Costa, among others, are labeled the 'positive-plus' logics. These logics lay the blame for trivialization on the classical theory of negation. They begin with an axiomatization of the positive (negation-free) fragment of classical logic, and then add to it a weakened account of negation, which blocks the derivation of arbitrary conclusions from $\{p, \neg p\}$, and even from $\{p \wedge \neg p\}$. The final (and preferred) place in the Priest/Routley/Norman taxonomy is reserved for relevance-based approaches. These include the more conservative American school of non-dialetheic relevance logic, and the relevance-based dialetheic school originating in Australia. These logics block trivialization by means of a substantial departure from classical logic, both with regard to how they treat negation and with regard to the conditional.

But the taxonomy I will apply here focuses instead on the various ways in which the classical consequence relations can be modified to avoid Triv. This *strategic* choice seems to me a better basis for distinguishing various approaches to paraconsistency. After all (as we shall see) giving a fair interpretation of the inferential tactics employed will depend on understanding the strategic maneuvers that lie behind those tactics. And the result is both comprehensive and finer grained. Any paraconsistent logic must change some part(s) of (3) and (4), just as every paraconsistent logic must make some tactical modification(s) of the axioms and rules of deduction. And the very same change in the tactical rules can be arrived at by quite different strategies for changing (3) and (4); a tactical taxonomy must simply pass over these differences.

The trivialization of inconsistency that arises from clauses (3) and (4) above has its roots in three distinguishable sources:

(A) The classical accounts of satisfiability and consistency

If we propose a logic according to which (some) classically unsatisfiable/inconsistent sets turn out to be satisfiable/consistent after all, then of course preserving this new form of satisfiability or consistency can be the basis of a consequence relation that does not trivialize all classically unsatisfiable or inconsistent sets of sentences. This raises a kind of puzzle, though not a terribly deep one, regarding whether such a logic is *really* paraconsistent, since it simply aims to preserve a new form of consistency, and attributes this new form of consistency to sets of sentences that classical logic regards as inconsistent.

(B) An unsatisfiable or inconsistent set lacks the only *property that the onsequence relation seeks to preserve. As a result, there are no grounds for constraining the consequences of such sets*

Thus the fact that the consequence relation is defined in terms of preserving consistency or satisfiability, rather than some other (desirable) feature of our premise set is essential to Triv. If we chose instead to preserve some new (desirable) feature of our premise sets, then the fact that a set lacks consistency or satisfiability need not imply that all constraints on the consequence relation are removed.

(C) The fact that the metalinguistic '$\rightarrow$' holds whenever the antecedent (i.e. that Γ' is a consistent or satisfiable extension of Γ) is false

If we were to alter our reading of this connective, we might create room to deny that clauses (3) and (4) hold trivially whenever Γ is inconsistent. This third strategy has a relevance flavor about it, though the paraconsistent relevance logics we will consider here all locate the problem with (3) and (4) in the classical account of consistency and satisfiability.

The roles of A, B and C in producing Triv lead to three basic strategies for avoiding it:

Strategy A: New accounts of 'truth' and 'consistency'

This is the road most traveled in paraconsistent logic. Its practitioners include those who propose a dialetheic account of truth. But they also include more conservative figures who take the new semantic values they propose for sentences to express epistemic commitment, or some other more metaphysically modest status than truth, *tout court*. On our taxonomy, any paraconsistent semantics that operates by non-trivially assigning *designated values* to all members of some classically unsatisfiable sets of sentences falls into this group. That is, this approach assumes that whenever α is not a (semantic) consequence of Γ in some logic, this must be because the semantics provides an acceptable valuation V such that for all $\gamma \in \Gamma$, $V(\gamma) \in \{v\text{: v is designated}\}$ and $V(\alpha) \notin \{v\text{: v is designated}\}$. Though truth is the standard example of a designated value, it's not necessary to interpret all, or even any, designated value as a formal theory of truth. N. D. Belnap, for instance, reads the values of Dunn's four-valued logic epistemically, as "told true," "told false," "told both" and "told neither." But even when we keep this interpretational latitude in mind, this account of the consequence relation is very constraining. It focuses all our attention on the assignment of values to sentences, the distinction between designated and undesignated values, and the consequence relation we get when $\Gamma \models \alpha$ is said to hold if and only if α is assigned a designated value whenever all the members of Γ are assigned designated values.

Strategy B: Preservationism

This approach has been less widely pursued. But it has, as we will see, some clear advantages over the first. The general idea has been put in various ways –

> Don't make things worse. (P. K. Schotch)
> Find something you like about your premises, and preserve it. (R. E. Jennings)

As we have seen, from the classical point of view, there is nothing worse than an inconsistent, unsatisfiable set of sentences. Classical logic aims only to preserve consistency and satifiability; once these are lost, there is nothing left that a classical logician cares to preserve. But there are other features of premise sets that are worth preserving. Non-triviality is the most obvious example, but we can (and will) be more specific. Going from a set of sentences that merely includes p and $\neg$p for some sentence p, to the set of all sentences clearly does make things worse. It takes us from a set that we could use to represent someone's commitments or the contents of an inconsistent theory in a non-trivial way, to a trivial representation of the commitments or the theory. And this suggests that we pursue precisely the constructive project that Jennings proposes.

This proposal creates a significant widening of the options before us. To propose another slogan, it liberates us from the tyranny of designated values. That is to say, unlike the first approach, it does not demand that we find a way to assign a designated value to the premises of a rejected consequence while assigning a non-designated value to the conclusion. In fact, the whole business of assigning values to sentences can be left just as classical logic has it, while we concern ourselves with features of sets that some inconsistent sets possess, and that are worth preserving.

Strategy C: A new metalinguistic '$\rightarrow$'

Classically, the '$\rightarrow$' in clauses (3) and (4) holds if either the antecedent is false, or the consequent is true. But if we demand some relevant connection between the antecedent and the consequent before declaring that conditions (3) and (4) are met, then the mere fact that the antecedent is false does not imply that the whole statement is true. Oddly, to the best of my knowledge those who have worked to apply relevance logic to these issues have all focused on the first approach to paraconsistency, rather than this one.

But there is an explanation of this. The tradition of relevance logic has endorsed the 'preserving designated values' account of consequence relations, while insisting that the consequence relation (and its object language reflection, the conditional) must respect considerations of relevance. Thus relevance logic demands that we reject both $\{p \wedge \neg p\} \models q$ as a consequence, and $(p \wedge \neg p) \rightarrow q$ as a conditional theorem. But if we continue to view consequences as a matter of preserving designated values, then to reject $\{p \wedge \neg p\} \models q$, we must be able to assign a designated value to $p \wedge \neg p$ while not assigning a designated value to q. And having done so, we will obviously be able to assign a designated value to the antecedent of our conditional, while not assigning a designated value to the consequent. So long as we treat the consequence relation as merely preserving designated values, there is no need to think of the corresponding conditional as preserving something more than designated values either. So from the perspective of traditional relevance logics, changing the metalinguistic $\rightarrow$ in the sort of way required here would involve doing just what the first approach to paraconsistency demands, namely producing a valuation which designates all the premises while failing to designate the conclusion.

However, a broader perspective might hold that $\rightarrow$ must preserve something other than designated values. This third approach is clearly distinct from the first, and worth pursuing independently. I recommend it to anyone interested in such preservationist conditionals.[8]

5 Paraconsistent Logics

Non-adjunctive logics

S. Jaskowski proposed the first formal paraconsistent logic in 1948 (reprinted in English in 1969). His approach was motivated by the idea of a discussion involving more than one participant, each contributing a consistent set of assertions. If we treat the assertions of each participant as 'holding' for the discussion, we end up with a potentially inconsistent set of sentences representing the overall product of the discussion. If we want a consequence relation under which we can reasonably close such sets of sentences, it must be a paraconsistent one. Moreover, the consequence relation should reject adjunction, that is the principle that $p,q \models p \wedge q$. After all, the fact that someone has contributed p to the discussion and someone has contributed q to the discussion in no way implies that anyone has contributed $p \wedge q$. Neither contributor need be in any way committed to $p \wedge q$. In fact, both contributors may regard p and q as incompatible with each other. Jaskowski's main concrete proposal for a logic that would respect these constraints is his D2.

To produce D2, Jaskowski appeals to the possibility operator of the strong modal logic S5. He lays his proposal out in terms of a correspondence between a discussive logic and an underlying modal logic. For each participant in the discussion, we consider the worlds satisfying all the claims they contribute to the discussion. Then we say that 'p' is *discussively* true if and only if '$\Diamond p$' is true at a world to which all and only these worlds are accessible. We say, further, that p discussively implies q if and only if $\Diamond p \rightarrow q$ is true in the underying modal logic, and (in an oddly asymmetrical definition) p and q are discussively equivalent if and only if $(\Diamond p \rightarrow q) \wedge (\Diamond q \rightarrow \Diamond p)$ holds in the underlying logic.

In S5, as in all standard modal logics, $\{\Diamond p, \Diamond q\} \not\models \Diamond(p \wedge q)$. Thus in D2, the corresponding discussive logic, $p,q \not\models (p \wedge q)$. So the rejection of adjunction is supported by this modal reading. Of course, the simplest view[9] today of how to represent a few people's commitments would be to have a separate accessibility relation for each individual, such that all and only the worlds at which someone's commitments are true are accessible to the actual world for that individual. The result, from the point of view of modal logic, would be a set of necessity operators, one for each individual, with the truth condition that p is discussively true if and only if $\Box_i p$ is true for some $\Box_i$. This approach to discussive logic would make it clear that (assuming for now the consistency of each individual's contribution to the discussion) each individual's contributions to the discussion should be closed under adjunction (in fact, will constitute a classical theory), even though the sum total of those contributions should not. Retaining some aggregative force here (particularly just how much we can or should retain) is an issue we will return to when we discuss the weakly aggregative logics of Schotch and Jennings.

This debate over the best means to arrive at a discussive logic aside, I'll finish here by making two points. First, Jaskowski's systems clearly fall within our first class of paraconsistent logics. p is 'true' in this logic if and only if '$\Diamond p$' true by the standards of the corresponding modal logic. The consequences of a set of sentences, Γ, are precisely those sentences α such that whenever $\Diamond\gamma$ is true for each γ in Γ, $\Diamond\alpha$ will also be true in the corresponding modal logic. Symbolically,

$\Gamma \models_d \alpha$ iff $\Diamond(\Gamma) \models_L \Diamond\alpha$, where $\models_L$ is the modal logic in question.

So preservation of truth for the admissible evaluations is the criterion of consequence for discussive logics. Second, the consequence relation here is the classical singleton consequence relation. A holds discussively if and only if some B such that {B} has A as a classical consequence has been added to the discussion by some participant. And the (propositional) theorems of the logic are simply the classical tautologies.

A different, very straightforward approach to non-adjunctive logic is due to Rescher and Brandom (1980). Beginning with the set of classical valuations, they propose two semantic operations: superposition and schematization. Applying superposition to two valuations produces a valuation assigning t to every sentence assigned t by either of the input valuations. Applying schematization produces a valuation assigning t to only those sentences assigned t by both of the input valuations. The full set of

Rescher–Brandom valuations results from closing the set of classical valuations under superposition and schematization.

The upshot is very similar to Jaskowski's D2: a logic whose theorems are just the classical tautologies, and whose consequence relation (as defined over the class of all such valuations) is the classical singleton consequence relation.

C-systems and weakened negation

Newton da Costa proposed his C-systems in the 1960s; a semantics for these logics did not emerge until later. But we will focus on the semantics here, since they are quite intuitive, and make it clear that da Costa's approach belongs solidly in the first category of our taxonomy. A da Costa evaluation maps every formula to t (the designated or 'true' value) or f, given an assignment to the sentence letters, as follows:

1. $v(A \wedge B) = t$ if and only if $v(A) = t$ and $v(q) = t$
2. $v(A \vee B) = t$ if and only if $v(A) = t$ or $v(B) = t$
3. $v(\sim A) = t$ if $v(A) = f$
4. $v(A) = t$ if $v(\sim\sim A) = t$

This class of valuations gives us the non-implicational fragment of the weakest C-system, C_ω. The C-systems are another example of our first class of paraconsistent logics: $\Gamma \models_{C\omega} \alpha$ if and only if every valuation assigning 1 to all members of Γ also assigns 1 to α. As is clear from the clauses for valuations, da Costa's logic is classical except in terms of how it treats the negation '~'. But (in part as an inevitable result of this choice) the negation is very nonclassical. In fact, Priest et al. (1989) argue that it is not a negation at all, but rather a "sub-contrary forming functor," that is, a functor *f* such that while $(p \wedge fp)$ can be true, $(p \vee fp)$ must be true. They point out as well that many very basic consequences involving classical negation fail for this negation. And while things get (for those wffs that behave 'consistently') more classical in the stronger C-systems, this does not help with the basic difficulties described by Priest, Routley, and Norman.

Further details of the C-systems and other systems that have emerged in the research programs of da Costa and his colleagues are left aside here for want of space.[10] The main point I want to emphasize for now is that the issue of how to tell a real negation from a pseudo-negation is a persistent problem for paraconsistent logic. It is by no means an easy question to answer. The initial assumption tends to be that classical negation is the paradigm case of a 'real' negation. But while this may represent the position a paraconsistent logician must *respond* to when arguing with critics of paraconsistency, it is unfair to begin by assuming that all the features of the classical negation are features that a good negation should have. Paraconsistent logic must insist that some, at least, of what classical logic does with negation is mistaken, at least in some applications.

Relevance logics

These logics have their roots in the program of relevance logic pioneered by Ackermann, and then developed and greatly extended by Anderson, Belnap, and their students. In these logics a tight correspondence is assumed to hold between conditional sentences and the consequence relation. But both are taken to be subject to strong constraints of relevance. In particular, a variable-sharing constraint is urged. $\{p, \neg p\} \models q$ is rejected on the grounds that there is no relevant connection of meaning between the premise set and the conclusion. Such a connection (at the propositional level) would demand at least that some variable appearing in the conclusion also appear in the premises. For reasons of space, we'll confine ourselves to examining three consequence relations that have their roots in the relevance program, avoiding the issue of conditional sentences whose logic and semantics would add quite a bit of complexity to the picture.

LP, the logic of paradox, is dramatically different from the C-systems in two respects. First, its semantics allows for sentences to be assigned both true and false at one and the same time. The values assigned to sentences in LP can be described as the nonempty subsets of the set {true, false}. Thus an assignment will assign one of the values {t}, {f}, or {t,f} to each sentence. We begin with an assignment of values to the sentence letters, and then extend it to the rest of the sentences following the clauses:[11]

1. (a) $t \in v(\neg A)$ iff $f \in v(A)$ (b) $f \in v(\neg A)$ iff $t \in v(A)$
2. (a) $t \in v(A \wedge B)$ iff $t \in v(A))$ and $t \in v(B)$ (b) $f \in v(A \wedge B)$ iff $f \in v(A)$ or $f \in v(B)$
3. (a) $t \in v(A \vee B)$ iff $t \in v(A)$ or $t \in v(B)$ (b) $f \in v(A \vee B)$ iff $f \in v(A)$ and $f \in v(B)$

The consequence relation is again defined in the standard way

$\Gamma \models_{LP} \alpha$ iff every such valuation making $t \in v(\gamma)$ for all $\gamma \in \Gamma$ also makes $t \in v(\alpha)$.

So once again we have an example of our first class of paraconsistent logics. The upshot here is quite clean and straightforward. The theorems (i.e. the consequences of the null set) are just the classical tautologies. And unlike the C-systems, the negation here looks very much like classical negation. The usual equivalences all hold:

$\{\neg A \wedge \neg B\} \models_{LP} \neg (A \vee B)$; $\{\neg(A \vee B)\} \models_{LP} \neg A \wedge \neg B$;
$\{\neg A \vee \neg B\} \models_{LP} \neg (A \wedge B)$; $\{\neg(A \wedge B)\} \models_{LP} \neg A \vee \neg B$
$\{A\} \models_{LP} \neg\neg A$; $\{\neg\neg A)\} \models_{LP} A$

Transitivity, as well as introduction and elimination inferences for $\wedge$ and $\vee$ are all retained.

The main classical principle that fails here (thereby preventing the trivialization of inconsistent premise sets) is disjunctive syllogism:

$A, (\neg A \vee B) \not\models B$

The failure of this principle is easy to see. If $v(A) = \{t,f\}$, while $v(B) = f$, then A has a designated value, as does $(\neg A \vee B)$, but B does not.

Having introduced truth-value 'gluts' here, that is sentences assigned both true and false, it might well seem natural to consider the possibility of truth-value gaps, that is sentences assigned neither true nor false. But in fact we can achieve the same effect with gluts alone, if we recognize an important fact about them: Sentences that are both true and false are both correctly assertable and correctly deniable.

Before we can apply this recognition to arrive at a different consequence relation in the relevance family, we will have to take a short detour back into classical logic, to bring clearly to mind some fundamental symmetries that apply there, and that have been lost in the transition to LP. Just as (for classical logic) 'truth' is the value that sustains the assertion of a sentence, that is that makes its assertion correct, 'false' is the value that sustains the denial of a sentence, that is that makes it correct to deny that sentence.

The notion that there are two distinct attitudes we can take with regard to declarative statements, assertion and denial, has often been rejected in favor of the view that we can make do with assertion alone. And in the context of classical logic there seems to be little reason to object to this. The position I take by denying that I am hungry seems equivalent in every important respect to the one I take by asserting that I am not hungry. However, a certain amount of expressive power is lost when we dispense with denial as a separate attitude, and with it we lose the capacity to express an important constraint on the consequence relation.

So far we have represented the consequence relation as a relation between sets of sentences on the left and individual sentences on the right. This is pretty standard practice, but it has some drawbacks. It makes the consequence relation something asymmetrical from the outset. But there are important and illuminating symmetries hidden behind this asymmetrical veil. We can reveal them by adopting a picture of the consequence relation that puts sets of sentences on both sides. We will say that $\Gamma \models_c \Delta$ iff every classical valuation satisfying all of Γ's members also satisfies some member of Δ. But now we can also say, equivalently, that this holds iff every classical valuation making all of Δ's members false also makes some member of Γ false.

If we have the notion of denial in hand, as well as the notion of assertion, then we can describe the condition under which $\Gamma \models_c \Delta$ holds in a different, but equivalent way. We can require that if every member of Δ is correctly denied in some valuation, then some member of Γ must also be correctly denied on that valuation. Of course, this is just the contrapositive of the truth-preserving account (i.e. correct assertability) of validity. However, preserving correct deniability from the right side of $\models$ to the left, as well as truth from the left to the right, can impose a real additional constraint on the consequence relation when our logic is not classical.

In particular, consider LP. What values shall we preserve from right to left, given that we preserve {t,f} and {t} from left to right? If the value {t,f} (often called 'both') is truly paradoxical,[12] one way to understand what we mean by that is to say that it sustains both the correctness of asserting a sentence that has it, and the correctness of denying that sentence. A sentence that is true and false both is both correctly assertable and correctly deniable. LP is the logic we get when we take the first point (that such sentences are correctly assertable) and ignore the second. As a result, LP preserves only {f} from right to left. But what we really should demand is that LP preserve both {f} and {t,f} from right to left:

$\Gamma \models_{LP^*} \Delta$ iff every LP valuation making $t \in v(\gamma)$ for all $\gamma \in \Gamma$ also makes $t \in v(\delta)$ for some $\delta \in \Delta$,

> and every LP valuation making $f \in v(\delta)$ for all $\delta \in \Delta$ also makes $f \in v(\gamma)$ for some $\gamma \in \Gamma$.

Such a logic preserves both LP-correct assertability from left to right, and LP-correct deniability from right to left.

What is SLP, this symmetrical version of the LP logic? The answer is, it's a step along the way to a logic familiar to students of relevance logic, namely first degree entailment (FDE).[13] The symmetrical form of LP behaves just like FDE except when the premise set cannot be consistently asserted *and* the conclusion set cannot be consistently denied. In these doubly (classically) trivial cases, this symmetrical version of LP trivializes just as classical logic does. One further step is required to arrive at FDE. We must coordinate our use of LP assignments to render the premises consistently assertable and the conclusions consistently deniable. If and only if there is an LP valuation 'satisfying' the premise set and an LP valuation 'falsifying' the conclusion set such that the two agree on their classical sub-valuations, and don't overlap on the sentence letters assigned 'both,' then the FDE consequence relation fails to hold between the premises and the conclusion.

The main point that I want to make about SLP and FDE here is that they very simply restore some fundamental symmetries present in classical logic, and given up in LP. For instance, no sentence is trivial on the left in LP, that is, there is no sentence A such that $\{A\} \models_{LP} \Delta$, for all sets Δ. But every sentence that is trivial on the right in classical logic is also trivial on the right in LP. These right-trivial sentences are, of course, the classical tautologies. That is, A is a classical tautology iff $\Gamma \models_{LP} A$, for all Γ. But in classical logic there is a perfect symmetry (duality) between the sentences that are trivial on the left and the sentences that are trivial on the right. For instance, $(p \wedge \neg p)$ is trivial on the left, and $(p \vee \neg p)$ is trivial on the right. LP forces us to surrender this symmetry; SLP and FDE restore it.

One final point is in order here. It is possible to get exactly the same consequence relations we have arrived at with the help of LP valuations while retaining a *fully classical* semantics. The trick is to change what the consequence relation is required to preserve. Rather than preserve truth from left to right (and falsehood from right to left), we can preserve a class of projections capable of producing a consistent image of our inconsistent set. We omit the details for reasons of space – but the general lesson is well worth drawing: one and the same consequence relation can be underwritten by quite different semantics. As a result, a paraconsistent consequence relation that is arrived at by one sort of change to our clauses (3) and (4) can (at least sometimes) also be achieved by a different sort of change.

Adaptive logics

Diderik Batens, inspired initially by L. Apostel, is the central figure in a research program working on a range of paraconsistent logics at the University of Ghent. Together with students and colleagues he has focused on a class of logics which he calls adaptive logics. The motives that lie behind these logics are a good fit with the preservationist approach to paraconsistent logic: Batens and his co-workers have been concerned with not making things worse when inconsistency rears its ugly head. However, the means by which they achieve these goals still focus, in the conventional way, on the preservation of designated values. Thus Batens says,

> An adaptive logic La localizes the abnormal properties of Γ, safeguards the theory from triviality by preventing specific rules of L (the initial, non-paraconsistent logic) from being applied to abnormal consequences of Γ, but behaves exactly like L for all other consequences of Γ . . . The (dynamic) proof theory of adaptive logics is based on the idea that a formula is considered to behave normally 'unless and until proved otherwise'. The semantics is better understood by another metaphor: La interprets Γ by eliminating its unnecessarily inconsistent L-models. For ACLuN2, e.g., the La-semantic consequences of Γ are the formulas true in the Lf-models of Γ that are minimally abnormal (not more inconsistent than required to make Γ true).[14]

Both the idea that we should not "make things worse," as Peter Schotch urges, and the idea that to constrain the consequences of a set of sentences we must find a way to make the set 'true' in some sense, are in the air here. As a result, these systems constitute a borderline case for this taxonomy. For reasons of space we will focus on the propositional fragment **PI** of the base paraconsistent logic, CLuN, and then explain briefly how the adaptive logic based on CLuN works.

PI includes sentence letters, and the usual truth-functional connectives. Their treatment is modified, however, from the familiar classical one, by the explicit surrender of the consistency assumption:

If, on some admissible valuation v, $v(A) = t$, then $v(\neg A) = f$.

This assumption is avoided by the simple expedient of making an assignment to the negations (the formulae that have '¬' as their main connective) a separate part of producing a valuation:

So we take as the base of our evaluation both an assignment v_b to the sentence letters and an assigment v_b to the negations:

$S \rightarrow t,f$
$N \rightarrow t,f$

The valuation v that results from this assignment is arrived at by the following rules

$v(S) = t$ iff $v_b(S) = t$ where S is a sentence letter
$v(\neg A) = t$ iff $v(A) = f$ or $v_b(\neg A) = t$
$v(A \wedge B) = t$ iff $v(A) = t$ and $v(B) = t$
$v(A \vee B) = t$ iff $v(A) = t$ or $v(B) = t$

This makes the main features of how this logic copes with inconsistency pretty clear. Note in particular that the effect of assigning values directly to the negations is only to make some negations true which would otherwise be false. Whenever a negation would be true given the initial assignment to the sentence letters alone, it remains true after the effect of the v_b assignment to the negations is factored in, since *either* the usual truth condition for '¬' or setting $v_b(\neg A) = t$ is sufficient to make $v(\neg A) = t$. So we can arbitrarily force any negation we like to receive the value true simply by assigning it true in v_b, but we cannot arbitrarily force negations to be false. Any negation whose truth follows from the classical components of the PI valuation will be true in the PI valuation. But in general many other negations will also be true.

It is dead simple, of course, to construct valuations that make an inconsistent set such as {p,¬p} true, while avoiding trivialization for some sets of sentences. The resulting logic is clearly paraconsistent.

PI and CLuN are clearly paraconsistent logics in the traditional, 'truth'-preserving mode. But the adaptive logics based on CLuN have something of the spirit of preservationism about them. A central idea for the propositional adaptive logic is the following theorem:

$$\vdash_c A \text{ iff, for some } C_1, \ldots C_n\ (n \geq 0),\ \vdash_{\mathbf{PI}} ((C_1 \wedge \neg C_1) \vee \ldots \vee (C_n \wedge \neg C_n) \vee A$$

That is, if A is a theorem of classical logic, then A will hold for **PI** as well, *unless* some of the C_i behave inconsistently. This theorem suggests a plan for the adaptive logic.[15] If we begin with the assumption that our premises are consistent, we can prove anything that classical logic allows us to prove. Suppose that we have proved A from our premises, using classical logic. Then by the theorem, there is some set $\{C_i\}$ of sentences whose consistent behavior is sufficient to assure us that A really does follow (in the **PI** sense) from our premises. But of course the assumption that the members of $\{C_i\}$ do behave consistently may turn out to be wrong. So the proof is *tentative*. It depends on the consistent behavior of the $\{C_i\}$ associated with A, which are kept track of at each step of the proof. If we should find, in the course of proving further consequences of our premises, that some C_i behaves inconsistently, we would have to *withdraw* our earlier proof of A. So proofs within an adaptive logic can require the deletion of earlier lines from the proof, based on what is shown later in the proof.

The result is a system of proof that allows us to derive the consequences that follow from our premises in the *minimally* inconsistent **PI** models of our premises. So these adaptive logics do involve a modification of the notion of a consequence. They make consequences turn on preserving something other than satisfiability. More inconsistent sets of sentences *are*, after all, still satisfiable, but we confine our attention to the minimally inconsistent sub-class of the models satisfying our premises, and define as a consequence what holds in these. From the point of view of our semantic clause (3), we have gone from requiring that a consequence α preserve the satisfiability of every satisfiable extension of our premises Γ to requiring that α preserve the minimally-inconsistent satisfiability of every minimally-inconsistent satisfiable extension of Γ. This is clearly a preservationist maneuver.

Weakly aggregative logics

With the work of P. K. Schotch and R. E. Jennings we finally arrive at a logic that is clearly and self-consciously preservationist. Rather than find a way of making premises true and conclusions false in some valuation in order to defeat an undesirable consequence, Schotch and Jennings identify a set of desirable properties that some inconsistent sets have, and propose a logic that preserves these properties.

To begin, we need the notion of a *level*, a generalization of consistency that our consequence relation will be required to preserve. The level of a set of formulae, Γ, is the minimum number, n, such that Γ can be partitioned into n consistent subsets. A formula α is a level preserving consequence (LPC) of Γ (Γ forces α, or, more formally, $\Gamma \vdash_{LPC} \alpha$), iff α is a classical consequence of some cell in every partition of Γ amongst n sets. Three important consequences of this are immediately apparent. First, any set not including a contradiction will have some well-defined level (possibly an infinite level, if Γ is an infinite set). Second, no set with a well-defined level will have any contradiction as an LPC. Third, no set including a contradiction has a well-defined level, since no partition of such a set will have only consistent cells. We assign the 'level' ∞ to such sets. Since such sets lack the property (having a well-defined level) that these logics aim to preserve, their consequences are trivial for this logic.

This non-adjunctive system goes beyond the completely non-adjunctive approaches of Jaskowski, and of Brandom and Rescher, allowing a weakened form of aggregation.[16] Given a level of 2, a set closed under LPC will include the disjunction of pairwise conjunctions of all triples in the set; given a level of 3, the set will include the disjunction of the pairwise conjunctions of all quadruples in the set, and so on. These disjunctions of pairwise conjunctions capture fully the adjunctive strength of forcing: Where n is the level of Γ, closing under the classical consequences of singleton subsets of Γ together with the rule 2/n + 1, which allows us to infer from any n + 1 formulae the disjunction of all their pairwise conjunctions, is consistent and complete with respect to forcing.[17]

Unlike the relevant systems[18] Schotch and Jenning's non-adjunctive logic agrees with some important intuitions concerning inconsistent input and conjunctions: if a set of claims including 'p,' '~p' is closed under forcing – the Schotch/Jennings inference relation – ~(p ∧ ~p) is forced but (p ∧ ~p) is not. So a computer using forcing would not regard itself as having been told the conjunction was true in such a case. This seems the right answer for an appropriately conservative inference engine to give.

Priest and Sylvan have objected to this, claiming that the computer has indeed been told the conjunction is true by implication.[19] Priest and Sylvan's argument assumes that the fact that conjunction just *is* the connective which gives a truth when the two things it joins are true implies that if adjunction fails as a rule of inference, then the connective it fails for can't be conjunction. But a crucial assumption about inference underlies this objection, *viz*. that truth-preservation is sufficient for the correctness of an inference. If correctness requires more than just truth preservation, the failure of {A,B} to imply (A ∧ B) does not show the logic has a non-standard truth condition for '∧.' And this is precisely the case for LPC. LPC requires preservation of level as well as truth. Adjunction *is* truth preserving, on Schotch and Jennings' view – it fails to be a rule of the system only because it fails to preserve level.

An interesting pragmatic element in inference emerges for non-adjunctive logics. Some individual conjunction-introductions are guaranteed to be level-preserving, but allowing conjunction-introduction in general leads to explosion. We must *decide* which conjunctions to adopt (if any). Our reasons for choosing some rather than others will normally derive from our epistemic aims. These often make some conjunctions indispensable; but they usually leave us also with a wide field of potentially interesting or desirable conjunctions whose value remains to be determined. On the question of which non-level increasing conjunctions to adopt and which to avoid, the logic is silent: adding them is just like extending any classical theory by adding further sentences consistent with, but not implied by, the theory. Which non-level increasing conjunctions we will accept depends on which are valuable – which seem required for effective application of the theory, which promise to produce interesting predictions without absurdity, and so on.

With regard to applications, this feature of LPC suits Bohr's theory of the hydrogen atom very nicely. The development of old quantum theory (OQT) involved a gradual clarification of which classical results can be applied in the quantum domain, and when. These results emerged from a process of investigating which classical results can be conjoined with quantum principles to good effect, in effect a choice of level-preserving conjunctions from among many candidates. Thus one prediction a forcing-based model makes concerning OQT is confirmed by the history of OQT: the addition of conjunctions of classical principles with quantum principles to the theory is ampliative, and requires independent theoretical and/or empirical justification. Conjunction introduction is not a trivial inference, but a substantial step that carries both risks and potential rewards.

6 Current Issues

Paraconsistent logic remains controversial. Many logicians still defend Triv and reject the entire field of paraconsistency as misguided. And even within the paraconsistent camp, the different approaches involve very different views of negation, consequence relations, and the nature of logic in general. As a result, proponents of one approach are often very critical of others. The breadth and range of positions in this field, only briefly and partially outlined here, makes giving a simultaneously brief and fair summary of the state of the field an impossible task. In this last section I want to briefly touch on two issues in paraconsistency that are at the center of my present work, and indicate how I hope, in further work, to provide some insight into them. The positions I will be sketching here are my own, and others will have their own responses (and, no doubt, trenchant criticisms of mine). So at this point I surrender all pretence of giving a balanced discussion.

Negation

As we have seen already, debates about negation, and especially about how to tell whether an operator in some logic is *really* a negation, have played a central role in the development of paraconsistent logic. On this issue I have a modest proposal: negation is *denial* in the object language. This is, I think, at least as credible a position as the standard relevance view that the object language conditional must express the (metalinguistic) consequence relation. On this view, a satisfactory paraconsistent logic needs an extended understanding of deniability that corresponds to the extended notion of assertability proposed by the logic. Only then can the symmetries between assertion and denial, negated and un-negated sentences, and the duality of $\wedge$ and $\vee$ be preserved. This leads us to a requirement on a satisfactory paraconsistent logic that some present systems meet and some fail: contraposition for the consequence relation:

$\Gamma \vdash \Delta$ iff $\neg(\Delta) \vdash \neg(\Gamma)$ $\quad (\vdash/\neg)$
($\neg(\Delta)$ is the set of sentences that results by negating each element of Δ.)

The idea here is that when we put a set of sentences in premise position, we are treating them as *in some sense assertable*. We take $\Gamma \vdash \Delta$ to say that if Γ is, in that sense, assertable, then so is some member of Δ. More explicitly, for every assertable extension of Γ, some element of Δ will be an assertable extension of the extension. So committing ourselves to Δ (in the sense of committing ourselves to accepting some member of Δ) begs no questions. It does not in any way extend the commitment we have already made in accepting Γ. But symmetrically, Γ should preserve the *deniability* of Δ, that is some member of Γ must be an acceptably deniable extension of every acceptably deniable extension of Δ. So if '$\neg$' really is the object language 'image' of denial, then from the right-to-left preservation of deniability must follow the left-to-right preservation of the assertability of the negations.

Interpretation

B. H. Slater has offered a Quinean objection to paraconsistent logic.[20] His claim is that if, in some logic, $\{A, \neg A\} \nvdash B$, this is just evidence (and well-nigh conclusive evidence) that '$\neg$' just isn't negation. In reply to this, I have pointed out that some preservationist logics, such as Schotch and Jennings' forcing, retain a fully classical semantics. They obtain $\{A, \neg A\} \nvdash B$ not by making A and $\neg$A somehow 'satisfiable' or 'consistent', but by finding another desirable property that $\{A,\neg A\}$ has, and that $\{A,\neg A,B\}$ lacks. But, taking this reply a step further, I have also shown that LP can be given a preservationist semantics based on using ambiguity to project consistent images of inconsistent sets of sentences. Details aside, the first lesson I think we should draw from this is that we should keep a healthy distance between our interest in various paraconsistent consequence relations and the particular semantics and definitions of consequence that we have used to arrive at them. The two are, in general, separable.

This also raises a more general question: can all non-preservationist paraconsistent logics be reinterpreted in preservationist terms? At least in some cases, such reinterpretations seem to be illuminating. And they do have the rhetorical advantage, when they begin from classical semantics, of saying nothing about the premise sets and the new consequence relation that classical logicians are inclined to deny. Preserving something other than satisfiability or consistency allows us not to argue with each other about what satisfiability or consistency *really* mean. Of course, this is not meant to cut off such discussions – as we remarked above, there is no reason to assume the classical account of these things must be the right one. But if our concern is to arrive at consequence relations for paraconsistent applications such as non-trivial inconsistent theories, we may do well to demonstrate the tenability and usefulness of such relations within a classical framework before (or even rather than) taking on the job of replacing it.

Notes

1 Strictly speaking, this should be set out more generally:

$$\frac{\Gamma \vdash \alpha,\ \Gamma \vdash \neg\alpha}{\Gamma \vdash \beta} \qquad \frac{\Gamma \models \alpha,\ \Gamma \models \neg\alpha}{\Gamma \models \beta}$$

2 Woods (1975: 165–7).

3 We can show that a set is inconsistent, by showing that some contradiction follows from it. And we can show that a set is consistent by presenting a model of it. But the process of finding a model of a set of sentences is not something we can go through systematically in such a way as to be sure that we will find a model at some finite point along the way if one is to be found. And when we fail to show a contradiction follows from some set of premises, our failure does not show that no contradiction follows.

4 Of course many have suggested that we reject excluded middle to avoid this unwelcome and radical conclusion. For example, we can add a 'gap' value to our semantics, which is neither true nor false, and assign 'gap' as the value of L. But as Graham Priest has often argued, this is not enough to resolve the problem. We can replace L with L′, the 'strengthened liar':

L′: This sentence is not true.

Now we can argue much as we did before. Suppose L′ is true. Then what it says must be true, so L′ is not true. Suppose L is not true. But this is precisely what L′ says of itself. So what L′ says is true, so L′ is true. Thus L′ is true if and only if L′ is not true. But now adding the value 'gap' (or some other, non-true value) to the usual values true and false is no help. See Priest (1995).

5 In fact, there are several notions of consistency, all closely related. A set that does not include both a sentence and its negation is called negation consistent; any set that doesn't include every sentence in the language is said to be absolutely consistent. In classical logic, the closure under deduction of any negation inconsistent set is the set of all sentences, i.e. absolutely inconsistent. In fact, the closure under deduction of any set whose closure under deduction is negation inconsistent is also absolutely inconsistent. The notion of consistency I am using here is that of any set whose closure under deduction is negation consistent. If we were sticking with classical logic, this would be equivalent to speaking of any set whose closure under deduction is not absolutely inconsistent.

6 Here we're using the notational convention that $\Gamma, \alpha = \Gamma \cup \{\alpha\}$. We could present (3) and (4) in slightly different form:

(3′) $\Gamma \models \alpha$ iff $\forall\Gamma'[(\Gamma'$ is a maximal satisfiable extension of $\Gamma) \rightarrow \alpha \in \Gamma']$.
(4′) $\Gamma \vdash \alpha$ iff $\forall\Gamma'[(\Gamma'$ is a maximal consistent extension of $\Gamma) \rightarrow \alpha \in \Gamma']$.

7 See Arruda (1989) for further details.

8 See Jennings and Johnston (1983), and D. Sarenac (2000) for work on such conditionals at the object-language level.

9 Based on the simplest approaches to epistemic logic.

10 See Arruda (1989) for more information on these systems and their applications.

11 See Kleene (1952) for an equivalent set of valuations, though Kleene does not treat the third ("paradoxical") value as designated.

12 See Priest (1995).
13 See Dunn (1986).
14 Batens (1998: 447).
15 For reasons of space the details must go unexplored here; see Batens (1998).
16 See Kyburg (1970).
17 See Schotch and Jennings (1989); Apostoli and Brown (1995).
18 See Belnap (1977) and Priest (1988).
19 Priest et al. (1989: 158).
20 Slater (1995).

References

Anderson, A. and Belnap, N. (ed.) (1975) *Entailment: The Logic of Relevance and Necessity.* Princeton, NJ: Princeton University Press.

Apostoli, L. and Brown, B. (1995) A solution to the completeness problem for weakly aggregative modal logics. *Journal of Symbolic Logic*, 60, 832–42.

Arruda, M. (1989) Aspects of the historical development of paraconsistent Logic. In G. Priest, R. Routley and J. Norman (eds.), *Paraconsistent Logic* (pp. 99–130). Munich: Philosophia Verlag.

Batens, D. (2000) A survey of inconsistency-adaptive logics. In Batens, Mortenson, Priest and Van Bendegem (eds.), *Frontiers of Paraconsistent Logic*. Baldock, Hertfordshire: Research Studies Press.

Batens, D. (1998) Inconsistency adaptive logics. In E. Orlowska (ed.), *Logic at Work: Essays Dedicated to the Memory of Helena Rasiowa*. Heidelberg: Springer.

Batens, D. (1989) Dynamic dialectical logics. In G. Priest, R. Routley and J. Norman (eds.), *Paraconsistent Logic*, Munich: Philosophia Verlag.

Batens, D., Mortenson, C., Priest, G., and Van Bendegem (eds.) (2000) *Frontiers of Paraconsistent Logic*. Baldock, Hertfordshire: Research Studies Press.

Belnap, N. D. (1977) How a computer should think. In Ryle (ed.), *Contemporary Aspects of Philosophy*. Oxford: Oriel Press; and also in Priest (1988) *Beyond Consistency*. Munich: Philosophia Verlag.

Béziau (2000) What is paraconsistent logic? In Batens, Mortenson, Priest and Van Bendegem (eds.), *Frontiers of Paraconsistent Logic*. Baldock, Hertfordshire: Research Studies Press.

Bohr (1913) On the constitution of atoms and molecules. *Philosophical Magazine*, 6, 1–25.

Brandom, D. and Rescher, N. (1980) *The Logic of Inconsistency*. Oxford: Basil Blackwell.

Braybrooke, P., Brown, B., and Schotch, P. K. (1995) *Logic on the Track of Social Change*. Oxford: Oxford University Press.

Brown, B. (1989) How to be realistic about inconsistency in science. *Studies in the History and Philosophy of Science*, 21, 2.

Brown, B. (1992) Rational inconsistency and reasoning. *Informal Logic*, XIV, 5–10.

Brown, B. (1993) Old quantum theory: a paraconsistent approach. *PSA 1992*, vol. 2, 397–411.

Brown, B. (1999) Yes, Virginia, there really are paraconsistent logics. *Journal of Philosophical Logic*, 28, 489–500.

Brown, B. (2000) Simple natural deduction for weakly aggregative paraconsistent logics. In Batens, Mortenson, Priest and Van Bendegem (eds.), *Frontiers of Paraconsistent Logic*. Baldock, Hertfordshire: Research Studies Press.

Brown, B. and Schotch, P. K. (1999) Logic and aggregation. *Journal of Philosophical Logic*, 28, 265–87.

da Costa, N. (1974) On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, XV, 497–510.

da Costa, N. and Alves (1977) A semantical analysis of the calculi C_n. *Notre Dame Journal of Formal Logic*, XVIII, 621–30.

Dunn, J. M. (1986) Relevance logic. In Gabbay (ed.), *Handbook of Philosophical Logic*. Dordrecht: Reidel.

Jaskowski, M. (1969) Propositional calculus for contradictory deductive systems. *Studia Logica*, 24, 143–57 (originally published in Polish in 1948, *Studia Scientarium Torunensis*, Sec. A II, 55–77).

Jennings, R. E. and Schotch, P. K. (1981) Some remarks on (weakly) weak modal logics. *Notre Dame Journal of Formal Logic*, 22, 309–14.

Jennings, R. E. and Johnston, D. (1983) Paradox-tolerant logic. *Logique et Analyse*, 26, 291–308.

Jennings, R. E. and Schotch, P. K. (1984) The preservation of coherence. *Studia Logica*, 43, 89–106.

Jennings, R. E., Schotch, P. K. and Johnston, D. (1980) Universal first order definability in modal logic. *Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik*, 26, 327–30.

Jennings, R. E., Schotch, P. K. and Johnston, D. (1981) The n-adic first order undefinability of the Geach formula. *Notre Dame Journal of Formal Logic*, 22, 375–78.

Johnston, D. (1978) A generalized relational semantics for modal logic, MA thesis, Department of Philosophy, Simon Fraser University, Barnaby, British Columbia.

Kleene, S. C. (1952) *Introduction to Metamathematics*. New York: Van Nostrand.

Kyburg, H. (1961) *Probability and the Logic of Rational Belief*. Middleton, CT: Wesleyan.

Kyburg, H. (1970) Conjunctivitis. In Marshall Swain (ed.), *Induction, Acceptance, and Rational Belief* (pp. 55–82). Dordrecht: Reidel.

Kyburg, H. (1997) The rule of adjunction and reasonable inference. *Journal of Philosophy*, XCIV, 109–25.

Kyburg, H. (1988) Full belief. *Theory and Decision*, 25, 137–62.

Priest, G. (2000) Motivations for paraconsistency: the slippery slope from classical logic to dialetheism. In Batens, Mortenson, Priest and Van Bendegem (eds.), *Frontiers of Paraconsistent Logic*. Baldock, Hertfordshire: Research Studies Press.

Priest, G. (1995) Reply to Parsons. *Canadian Journal of Philosophy*, 25.

Priest, G. (1988) *Beyond Consistency*. Munich: Philosophia Verlag.

Priest, G., Routley, R., and Norman, J. (eds.) (1989) *Paraconsistent Logic: Essays on the Inconsistent*. Munich: Philosophia Verlag.

Priest, G. and Tanaka (1998) Paraconsistent logic. In Stanford's online philosophy encyclopedia, http://plato.stanford.edu/entries/logic-paraconsistent/logic-paraconsistent.html.

Sarenac, D. (2000) The preservation of meta-valuational properties and the meta-valuational properties of implication. In Woods and Brown (eds.), *Logical Consequence: Rival Approaches*. London: Hermes.

Schotch, P. K. and Jennings, R. E. (1980a) Inference and necessity. *Journal of Philosophical Logic*, 9, 327–40.

Schotch, P. K. and Jennings, R. E. (1980b) Modal logic and the theory of modal aggregation. *Philosophia*, 9, 265–78.

Schotch, P. K. and Jennings, R. E. (1989) On Detonting. In G. Priest, R. Routley and J. Norman (eds.), *Paraconsistent Logic* (pp. 306–27). Munich: Philosophia Verlag.

Slater, B. H. (1995) Paraconsistent logics? *Journal of Philosophical Logic*, 24, 451–4.

Woods, J. (2000) Pluralism about logical consequence: resolving conflict in logical theory. In Woods and Brown (eds.), *Logical Consequence: Rival Approaches*. London: Hermes.

Woods, J. and Brown, B. (eds.) (2000) *Logical Consequence: Rival Approaches*. London: Hermes.