When inconsistency is inescapable: A survey of paraconsistent logics

Graham Priest
Department of Philosophy, University of Western Australia, Nedlands, Western Australia 6009, Australia

Received July 1987

A logic is *paraconsistent* iff the inference $\{\alpha, \neg\alpha\} \vDash \beta$ fails in the logic in general. In this article I will explain various paraconsistent logics and some of their applications and ramifications. There are at least four different approaches to paraconsistent logic: (i) The non-adjunctive approach (due originally to Jaskowski). Characteristically this rejects the inference $\{\alpha, \beta\} \vDash \alpha\wedge\beta$. (ii) The positive-plus approach (largely due to da Costa). This adds to standard positive logics a non-truth-functional negation. (iii) The relatedness approach (due originally to Smiley). Characteristically, this rejects the transitivity of deducibility. (iv) The De Morgan approach (due originally to Anderson and Belnap). Characteristically, this approach gives up the disjunctive syllogism, $\{\alpha\wedge(\neg\alpha\vee\beta)\} \vDash \beta$. The importance of paraconsistent logics is that they allow for the recognition of the existence of theories that are inconsistent but non-trivial, and consequently for the investigation of these. Such theories may occur in many domains: (i) *Automated reasoning*. Any sophisticated data base is liable to be inconsistent due to multiple sources, the undecidability of inconsistency, etc. Moreover, any reasonably powerful AI reasoning system is liable to end in inconsistency, due to semantic paradoxes. (ii) *Semantics*. Any adequate theory of meaning must, to avoid self-refutation, be theoretically capable of giving the semantics of the language in which it, itself, is expressed. Such theories characteristically end up in inconsistency due to the semantic paradoxes. (iii) *Set theory*. A naive theory of sets based on the unrestricted comprehension axiom $\exists y \forall x (x\varepsilon y \leftrightarrow \varphi)$, though inconsistent, provides for many of the set-theoretic operations required in parts of mathematics, such as category theory, not available in theories such as ZF. The recognition of important inconsistent theories has many philosophical implications. One immediate one concerns Goedel's incompleteness theorem. Many have tried to make philosophical capital out of the claim that for any (suitably strong) theory there are true statements that are not provable. Such capital is rendered worthless by paraconsistent logics. For Goedel's theorem applies only to *consistent* theories. Yet many global theories are inconsistent as (i) – (iii) above illustrate. Challenging and important new theories are fairly rare in logic, but as, I hope, should be clear, paraconsistent logic is just such a theory.

'n Logika is *parakonsistent* as en slegs as die afleiding $\{\alpha, \neg\alpha\} \vDash \beta$ in dié logika in die algemeen faal. In hierdie artikel verduidelik die outeur verskeie parakonsistente logikas en van hulle toepassings en vertakkings. Daar is ten minste vier verskillende benaderings tot parakonsistente logika: (i) die saamvoegingslose benadering (oorspronklik te danke aan Jaskowski). Kenmerkend verwerp hierdie benadering die afleiding $\{\alpha, \beta\} \vDash \alpha \wedge \beta$. (ii) Die positief-plus-benadering (in 'n groot mate te danke aan Da Costa). Dit voeg aan standaard-positiewe logikas 'n nie-waarheidsfunksionele negasie toe. (iii) Die verwantskap-benadering (oorspronklik aan Smiley te danke). Kenmerkend hiervan is die verwerping van die transitiwiteit van deduseerbaarheid. (iv) Die De Morgan-benadering (oorspronklik te danke aan Anderson en Belnap). Kenmerkend hiervan is die prysgawe van die disjunktiewe sillogisme, $\{\alpha\wedge(\neg\alpha\vee\beta)\} \vDash \beta$. Die belangrikheid van parakonsistente logikas is dat hulle die moontlikheid laat vir die erkenning van die bestaan van teorieë wat inkonsistent maar nie triviaal is nie, en gevolglik vir die ondersoek hiervan. Sodanige teorieë kan op baie terreine voorkom: (i) *Geoutomatiseerde redenering*. Enige gesofistikeerde databasis kan inkonsistent wees as gevolg van 'n verskeidenheid bronne, die onbeslisbaarheid van strydigheid, ens. Bowendien kan enige redelik kragtige kunsmatige intelligensie-stelsel vir redenering op strydighede uitloop vanweë semantiese paradokse. (ii) *Semantiek*. Enige adekwate betekenisteorie moet, ten einde selfweerlegging vry te spring, teoreties in staat wees om rekenskap te gee van die semantiek van die taal waarin die teorie self geformuleer is. Kenmerkend van sodanige teorieë is egter dat hulle vanweë die semantiese paradokse op strydighede uitloop. (iii) *Versamelingsteorie*. 'n Naïewe teorie van versamelings gebaseer op die onbeperkte aksioma van versamelingsvorming, $\exists y \forall x (x\varepsilon y \leftrightarrow \varphi)$, voorsien, hoewel dit inkonsistent is, talle van die versamelingsteoretiese operasies wat in dele van die wiskunde, soos kategorieteorie, benodig word, maar wat nie voorhande is in teorieë soos ZF nie. Die erkenning van belangrike inkonsistente teorieë het baie filosofiese implikasies. 'n Voor-die-hand-liggende een raak Gödel se onvolledigheidsteorema. Baie het filosofiese munt probeer slaan uit die aanspraak dat ten opsigte van enige (gepas sterk) teorie daar waar bewerings is wat nie bewysbaar is nie. Parakonsistente teorieë verydel dié hoop op filosofiese wins, want Gödel se teorema geld slegs *konsistente* teorieë. Baie totaliteitsteorieë is egter inkonsistent, soos punte (i) – (iii) hierbo aantoon. Uitdagende en belangrike nuwe teorieë is redelik skaars in logika, maar die outeur hoop dat dit duidelik sal wees dat parakonsistente logika juis so 'n teorie is.

Introduction[1]

Paraconsistent logic is a relatively new branch of logic; its history is barely 35 years old. Yet already it has shown itself to be an important development in modern logic. In this article I will try to give some idea of the content and importance of the subject, assuming that the reader knows little about it. Because the subject is far too large to give even a comprehensive survey in a single article, I shall not attempt this. Rather, I shall give a selective survey which, I hope, conveys the spirit of the subject. For the same reason I shall not give any detailed proofs in this article. The proofs of facts cited are either sufficiently elementary to be left to the reader, or else referenced.[2]

Let us start with a definition of the subject. Suppose that $\vdash$ is the consequence relation of a logic. Let us say that it (and, derivatively, the logic itself) is *explosive* if the inference: $\{\alpha, \neg\alpha\} \vdash \beta$, *ex contradictione quodlibet (ECQ)*, is valid. Classical logic, intuitionist logic, and most other logics commonly met are explosive. A logic is *paraconsistent* if it is not explosive. The importance of paraconsistent logic lies precisely in the fact that it allows for the existence of inconsistent but non-trivial theories; that is, it allows us to reason in inconsistent situations without our conclusions exploding to totality. This is important since there are many situations which are perforce inconsistent, yet where we still need to discriminate in the conclusions we draw. I will return to this in the second half of the article. In the first part I will describe the various approaches to formal paraconsistent logic.

1. Approach to paraconsistency

Since the defining characteristic of paraconsistent logics is the failure of the inference *ECQ*, a good place to start a consideration of the subject is with C.I. Lewis's well-known argument for the principle (Lewis & Langford 1959: 250). The argument (slightly modified for present purposes) goes as follows:

(1) α	Assumption
(2) $\neg\alpha$	Assumption
(3) $\neg(\alpha\wedge\neg\beta)$	From (2)
(4) $\alpha\wedge\neg(\alpha\wedge\neg\beta)$	From (1) and (3)
(5) β	From (4)

This argument depends on four principles. The first is the principle $\{\neg\varphi\}\vdash\neg(\varphi\wedge\psi)$ employed at line (3). The second is adjunction principle $\{\varphi, \psi\} \vdash \varphi\wedge\psi$, employed at line (4). The third is the disjunctive syllogism, $\{\alpha\wedge\neg(\alpha\wedge\neg\beta)\} \vdash \beta$, employed at line (5). The final principle is not explicit at any line but is implicit in the notation; this is the transitivity of deducibility. Given that (5) follows from things that follow from (1) and (2), it, too, follows from (1) and (2). The rejection of each of these four principles provides an (eco)logical niche for a family of paraconsistent logics; and like most such niches, all four are inhabited, as I will now show. In what follows, I intend to discuss only propositional logics. This is not because quantificational extensions of the logics I shall mention are problematic. It is because, as the above argument indicates, the real issues involved here concern principles to which quantification *per se* is irrelevant. In fact, each logic in the following four families can be extended to cope with quantification in simple and obvious ways. The matter may therefore be safely left to the cogitations of the reader.

2. Non-adjunctive logics

I shall call logics which reject the adjunction principle, naturally enough, non-adjunctive logics. Such logics were the first of the four kinds of paraconsistent logics to be investigated; investigations were initiated by the Polish logician Jaskowski in 1948 (cf. Jaskowski, 1969). Jaskowski's idea was simple: suppose that the information we have is provided by a number of different sources, each internally consistent but each possibly conflicting with others. Then we shall not necessarily wish to combine information from different sources to make inferences. Consequently, multi-premise inferences, such as adjunction, are suspect.

Jaskowski made his approach rigorous essentially as follows. Suppose we have a Kripke interpretation, $\mathfrak{A}$, for the modal logic S5. Then we may identify the information, or discourse, of each source with what is true in some one possible world in $\mathfrak{A}$. Let us therefore define: α is *discursively true* in $\mathfrak{A}$ iff $M\alpha$ is true (at any world) in $\mathfrak{A}$. An inference is said to be *discursively valid* iff it preserves discursive truth in every interpretation. I leave it as a trivial exercise to show that both adjunction and *ECQ* are discursively invalid.

Less happily, perhaps, it is also simple to show that the inference $\{\alpha, \alpha\supset\beta\} \vdash \beta$ is not discursively valid. Since the existence of an implication connective satisfying *modus ponens* is obviously desirable Jaskowski was forced to find one. The one he favoured, discursive implication, $\supset_d$, is defined as follows: $\alpha\supset_d\beta$ is $M\alpha\supset\beta$. As may be checked, if α and $\alpha\supset_d\beta$ are both discursively true, so is β. Indeed, discursive implication behaves very much like material implication. In particular, as Jaskowski proved, the discursive logical truths containing $\supset_d$ as the sole connective are exactly the classical tautologies.[3]

Jaskowski's original construction may be varied to give a whole family of logics.[4] Obvious variations involve changing the base modal logic and/or taking a modality other than M in the definition of discursive truth. More sophisticated variations are also possible.[5] Typically in these logics, however, single-premise inference is quite orthodox. Thus $\{\alpha\wedge\neg\alpha\} \vdash \beta$ is valid. It follows then, that adjunction must fail if the logic is to remain paraconsistent. Conjunction must, therefore, behave non-standardly.

3. Positive-plus logics

Let us move on to the second logical niche isolated above, that where the principle $\{\neg\varphi\} \vdash\neg(\varphi\wedge\psi)$ fails. In fact, this failure is more profitably seen as a result of a more fundamental failure. If we are not to follow the path of discursive logic, but to have a normally behaving conjunction then we will have $\{\varphi\wedge\psi\} \vdash \varphi$. Thus, in this niche contraposition must fail; and this may be thought of as the fundamental principle rejected by this approach.

This line was suggested first by the Brazilian logician da Costa in 1963.[6] Da Costa, in fact, took it that not only conjunction should behave in a standard fashion, but that all positive connectives should so behave. His idea was therefore to obtain a paraconsistent logic by grafting a non-standard negation on to classical or intuitionist positive logics. This is why I call this approach 'positive-plus'.

The approach may be illustrated simply as follows.[7] Let a *positive-plus evaluation* be any map from the language of the classical propositional calculus to $\{0, 1\}$, which is truth functional in the normal way with respect to positive connectives, but is arbitrary and may be non-truth-functional on negation. Thus, the value of α under an evaluation may be 1 or 0 quite independently of the value of $\neg\alpha$ under the evaluation. *Positive-plus consequence* is defined, in the natural way, as truth preservation under positive-plus evaluations. As may easily be checked, both contraposition and *ECQ* are positive-plus invalid.

The logic just specified is not very interesting. For the conditions on negation are so weak that there are no principles of inference which concern negation essentially. To make the logic interesting, extra conditions must therefore be imposed on the evaluations of negations. A simple one (cf. Batens, 1980) is the condition that at least one of α and $\neg\alpha$ must be true under an evaluation. This validates the law of excluded middle. Further conditions generate members of da Costa's hierarchy of logics C_i, for natural number i. And if we start with valuational semantics, or Kripke semantics, for positive intuitionist logic and graft on a non-truth-functional negation in the same way, we obtain, amongst other things, da Costa's logic C_ω.

Whilst we can obtain many of the standard properties of negation by imposing extra conditions on the evaluation of negation, it should be noted that contraposition must always fail on this approach, at least if it is to remain paraconsistent.

For since the positive logic is standard (classical or intuitionistic), $\alpha \supset (\neg\beta \supset \alpha)$ is a logical truth. Hence if we had contraposition in the form, $(\neg\beta \supset \alpha) \supset (\neg\alpha \supset \beta)$, it would follow by the (standard) transitivity of $\supset$, that $\alpha \supset (\neg\alpha \supset \beta)$; whence the logic would become explosive.

4. Filter logics

Mention of the transitivity of implication brings us to another of the ecological niches I mentioned above: that where this is rejected. Formally, this approach was first proposed by Smiley in (1959). In a reasonably general form the approach is as follows. Let $F(\alpha, \beta)$ be some condition on pairs of formulas.[8] For reasons that will become clear, I will call this relation a *filter*. Let us say that an inference $\{\alpha\} \vdash \beta$ is *filter-prevalid* iff it is classically valid and $F(\alpha, \beta)$ holds. An inference is *filter-valid* iff it is a substitution instance of an inference that is filter-prevalid. (The complication concerning prevalidity may be necessary to ensure that the logic is closed under substitution).

In its most general setting, filter-validity is not a particularly interesting notion. By choosing the empty filter we can make every inference invalid. But Smiley, and those who have followed him in this approach, had a particular kind of filter in mind. The thought is that for an inference to be valid it should do more than just preserve truth; there must be some connection of relevance or meaning between premise and conclusion. The filter, F, was meant to disqualify pairs of formulas failing this connection. This still leaves plenty of formal scope as to what the filter should be. Smiley's original idea was to take the relation $F(\alpha, \beta)$ to be: α is not a contradiction and β is not a tautology. In this case, each of the individual inferences used in the argument of section 1 is filter-valid, but *ECQ* is obviously not. It follows, then, that transitivity fails.

By varying the formal filter, we can generate a whole family of filter logics.[9] Some of these filters will produce logics that are relevant in the technical sense. A logic is (technically) *relevant* iff whenever the inference $\{\alpha\} \vdash \beta$ is valid, according to the logic, α and β share a propositional variable. As a moment's thought shows, Smiley's filter above produces a relevant logic. More simply, and not equivalently (but with the same effects on transitivity and *ECQ*), we might just take the filter to be variable sharing itself. Such a filter-logic is obviously relevant.[10]

5. De Morgan logics

There is also a close connection between relevance and the fourth class of logics I mentioned in section 1, those where the inference disjunctive syllogism characteristically fails,[11] De Morgan logics. Some of these were first proposed seriously by the American logicians Anderson and Belnap (cf. 1975) in the early 1960s as an analysis, amongst other things, of the informal notion of relevance.

I call these logics De Morgan logics, not because of any intrinsic connection with Augustus, but because their algebraic semantics[12] centre on the notion of a De Morgan lattice (due largely to the American logician Dunn).[13] A *De Morgan lattice* is a distributive lattice with an involution operator, i.e., an operator *, satisfying the conditions i) $a = a^{**}$ and ii) if $a \le b$ then $b^* \le a^*$. A *De Morgan evaluation* for a propositional language containing the connectives $\wedge$, $\vee$ and $\neg$, is a map from formulas to a De Morgan lattice such that the meet, join and involution operators are the interpretations of conjunction, disjunction and negation respectively. An inference [14] $\{\alpha\} \vdash \beta$ is *De Morgan valid* iff for every De Morgan lattice and every evaluation on that lattice, v, $v(\alpha) \le v(\beta)$.

The simplest De Morgan lattice, other than the 2-element Boolean algebra, is the 3 element algebra {f, p, t}, (ascendingly) ordered as shown, where * maps t to f, vice versa, and p to itself. Taking the evaluation which maps α to p and β to f, we obtain counter-examples to both the disjunctive syllogism and *ECQ*.

The logic just specified is, in fact, Anderson and Belnap's (1975, ch.3, sec.18) logic of First Degree Entailment, and is, arguably, the most basic De Morgan logic. Other such logics are obtained by extending the above semantics in two ways. First, since De Morgan lattices are not guaranteed a maximal element, it is impossible to use such an element to define theoremhood in the way standard in algebraic semantics. To rectify this problem, we suppose that the lattice comes with a designated element, μ, such that theoremhood corresponds to taking an algebraic value greater than or equal to μ under any evaluation.

Secondly, and more importantly, to cope with languages that contain an implication connective, $\rightarrow$, we need to add a new binary operator, $\Rightarrow$, to the algebra. This is to be the semantic interpretation of the connective.[15] Putting various conditions on $\Rightarrow$ and μ (the most obvious of which is: if $a \le b$ then $\mu \le a \Rightarrow b$) produces a rich variety of De Morgan logics, which have been investigated by many people.[16] Amongst these are the higher degree logics of Anderson and Belnap, E and R. By no means all the logics generated are technically relevant, however. Indeed, some of the simpler and important ones are irrelevant.[17]

I have now reviewed the four families of paraconsistent logics: non-adjunctive logics, positive-plus logics, filter-logics, and De Morgan logics. As we have seen, they are not only distinct, but have been motivated by often quite different considerations. All, however, sustain the possibility of reasoning non-trivially in inconsistent situations. It is time to turn to some examples of situations where this is important. There are a number of these.

6. Automated reasoning

Let us start with the area of automated reasoning. One of the most obvious uses of computers is the storage of large amounts of information. But we require computers not only to store the information for us but to operate on it: to sort it, search it and, crucially, to infer from it. Now, as anyone who is minimally connected with data collection knows, data collected are liable to be inconsistent, both because of mistakes made in entering it and because of multiple sources. This may not be too much of a concern for a simple relational database, for such data-bases have no real way of expressing negation at all; but it is certainly a problem for more sophisticated data bases operating with theorem-provers. For if the theorem-prover implements an explosive logic (which, of course, most standard theorem-provers do) the automated reasoning may give us a totally arbitrary conclusion based on this inconsistency.

Neither is there any way we would automatically expect to pick this problem up. We cannot rely on the machine telling us that the data are inconsistent, since it might not discover this *en route* to its conclusion. We might be alerted to it by the fact that the computer told us 'yes' to every question we asked it. But this might not happen: the answers to some questions may not be found, either because the memory space is too small, or because the heuristic search used fails to detect them. And even if it, or we, did discover that the data was inconsistent, this would not be a great deal of help. Ideally we might like to get rid of the inconsistency in the data, but there is no algorithm for this. Perhaps, in the future, we will have heuristics for modifying inconsistent data-bases to get rid of a particular contradiction; but none exists at the moment. And even if we had such a heuristic, we are stuck with the inconsistent data until it is applied (which may take considerable time). Moreover, since it *is* only a heuristic, the data are still liable to be inconsistent afterwards. We therefore seem to be stuck with the problem of inconsistency here. The situation is tailor-made for paraconsistent logic.

The next question, then, is how to implement it effectively. As far as I know, the only class of paraconsistent logics for which this has been considered in any detail is the last one we considered: De Morgan logics. There are several neat algorithms for implementing the logic of first degree entailment. For example, the following, due to the Chinese computer scientist, Fangzhen Lin,[18] which uses semantic tableaux. Suppose that we draw up the truth-tree for a formula using the standard rules of classical logic (so that, for example, from a node of the form $\alpha \wedge \beta$ we descend to nodes α and β; and from a node of the form $\neg(\alpha \wedge \beta)$ we split the branch and descend to one node $\neg\alpha$ and one $\neg\beta$.[19] Given two branches of such trees let us call them *complementary* if there is an atomic formula which occurs negated in one branch and unnegated in the other. Then the inference $\{\alpha\} \vdash \beta$ is valid in the logic in question iff for every branch of the tree of α there is a complementary branch in the tree for $\neg\beta$ and vice versa. (Incidentally, this algorithm can be modified in a straightforward way to handle quantification.)

Once we consider De Morgan logics which contain an implication connective the situation is less satisfactory. For a start, mainly propositional logics have been investigated,[20] and even here the algorithms are quite complex. Perhaps the major problem is that the best understood and most efficient algorithm for implementing classical and similar logics involves the principle of inference called resolution: $\{\neg\alpha \vee \beta, \neg\beta \vee \gamma\} \vdash \neg\alpha \vee \gamma$. This is but a variant of the disjunctive syllogism, whose failure is, as we saw, the distinguishing mark of De Morgan logics.[21] Thus, the standard techniques are inapplicable here. It would be nice to have a method for De Morgan logics as simple and powerful as resolution is for classical logics; but none such is known. An extra complicating factor is that several of these *propositional* logics are known to be undecidable (though recursively enumerable).[22]

7. Semantics

Before we leave the topic of automated reasoning, let us note that there is another and, perhaps, more profound reason why this is forced into the inconsistent. To explain this I need to talk about truth for a moment. Truth is a predicate of statements, beliefs, or other cognitive entities, which we may, for the present, take to be represented by sentences. It is characterized, *prima facie* at least, by the principle known as the Tarski T-scheme: $T\underline{\alpha} \leftrightarrow \alpha'$ where α is any sentence, $\underline{\alpha}$ is its name, and α' states its meaning.[23] Now in many languages there is a variety of means (using demonstratives, definite descriptions, arithmetic diagonalization, etc.) to construct a sentence, β, that means that $\neg T\underline{\beta}$. (This very sentence is not true.) Substituting this in the T-scheme gives $T\underline{\beta} \leftrightarrow \neg T\underline{\beta}$, whence it follows that $T\underline{\beta} \wedge \neg T\underline{\beta}$.

This is the liar paradox, and is but one of a large number of paradoxes known to arise in connection with truth and related notions, such as satisfaction and definition. What has it to do with automated reasoning? Simply this: any artificial intelligence system capable of more than very limited application must be able to reason about its own cognitive states, those of others, and in particular, their truth or otherwise.[24] Thus the system will be inconsistent.

The major suggestions concerning how to avoid the semantic paradoxes are to the effect that the T-scheme itself must be weakened somehow. These suggestions are all notoriously problematic.[25] Moreover, the weakened forms of the T-scheme are not adequate to capture standard and unproblematic cognitive reasoning (cf. Priest, 198+). There is not time to go into this now. But it should be noted that this avenue of escape is ruled out in another context where paraconsistent logic finds another natural application: semantics.

Semantics is that branch of logic/linguistics which aims to spell out a theoretical understanding of meaning, both in general and of particular languages. How this should be done is still a matter of contention. But all the general accounts of meaning we have, insist that to spell out the meaning of an indicative sentence is, in some sense, to spell out its truth conditions.[26] The sentence which spells out the truth conditions of a sentence is, of course, the instance of the T-scheme for that sentence. Thus theoretical linguistics is committed to the T-scheme in some form or other, and hence to inconsistency.

The problem of inconsistency in linguistics has been avoided rather than faced, by and large. It has been avoided by the simple expedient of giving semantics for languages (or fragments of natural language) which do not themselves contain semantic notions. Once it is faced, it is clear that general semantics must be based on paraconsistent logic. In this way it is possible to construct theories of truth and meaning for languages in which the theory itself is couched (and which, *ipso facto*, contain the notion of truth), as has now been demonstrated.[27]

Before we leave the issue of semantic paradoxes, there is another important observation worth making. This concerns the principle of inference called absorption: $\{\alpha \rightarrow (\alpha \rightarrow \beta)\} \vdash \alpha \rightarrow \beta$. As was shown essentially by Curry in (1942), the T-scheme plus absorption (and *modus ponens*) leads to triviality quite independently of the deduction of contradiction.[28] The argument goes as follows: by self-reference of any form, we can construct a sentence, γ of the form $T\underline{\gamma} \rightarrow \beta$, where β is arbitrary. (If this sentence is true, β.) The T-scheme for the sentence is: $T\underline{\gamma} \leftrightarrow (T\underline{\gamma} \rightarrow \beta)$. Absorption, left to right, gives $T\underline{\gamma} \rightarrow \beta$; then *modus ponens*, right to left, gives $T\underline{\gamma}$; whence, again by *modus ponens*, β.

This observation is important since it shows, as might be expected anyway, that not all paraconsistent logics are suitable for all inconsistent situations. In particular, all non-adjunctive logics and positive-plus logics have absorption as a valid principle; most of the standard filters also fail to filter out absorption. Hence De Morgan logics are the only class of logics suitable for use in this context; and not all of these can be used either.[29]

8. Set theory

The final kind of inconsistent situation I wish to mention is set theory. This may appear somewhat surprising initially; for set theory is nowadays tacitly identified with Zermelo-Fraenkel set theory, or some other set theory pegged to the cumulative hierarchy. And such theories are (we all hope) consistent. But, of course, the semantic paradoxes of the last section are only one kind of paradox; and they have a close cousin in the set-theoretic paradoxes. Let us review the historical situation.

At the end of the last century set theory made its appearance as a fully fledged theory, and was even formalized by Frege. The axioms of set theory which seemed to Frege to capture our intuitive notion of set (and no one would ever have thought to challenge this view but for the paradoxes) amount to conditions for the existence and identity of sets, Abstraction and Extensionality:

$\forall y\ (y\varepsilon\{x;\varphi(x)\} \leftrightarrow \varphi(y)\ )$

$\forall x(\varphi(x) \leftrightarrow \psi\ (x)) \rightarrow \{x;\varphi(x)\}\ =\ \{x;\psi(x)\}$

where in the first of these y is free for x in φ.[30] As was discovered, given only minimal logical principles of inference, the axioms are inconsistent. The quickest contradiction is Russell's. Just take '$\neg x\varepsilon x$' for $\varphi(x)$ in the abstraction axiom, to get $r\varepsilon r \leftrightarrow \neg r\varepsilon r$ and hence $r\varepsilon r \wedge \neg r\varepsilon r$, where r is $\{x;\neg x\varepsilon x\}$.

Now, much effort was put into trying to find a consistent sub-theory of this theory, adequate to do justice to the notion of set. The result is Zermelo-Fraenkel set theory (or some near equivalent). And it must be admitted that it does pretty well, at least as long as we stick to local theories (i.e., theories concerning sets such as the reals, of some bounded rank). Once we move to global theories, however, it comes unstuck. This is particularly clear in the case of category theory.[31] Notoriously, (ZF) set theory is unable to provide the conceptual set theoretic wherewithal for this.[32] In particular, it is unable to form or operate on large categories, such as the category of all sets, of all groups, or of all categories.

This is a direct result of the fact that the set-existence axioms of ZF are but a pale shadow of those of Frege, which, as he thought, do capture the set theoretic constructions inherent in thought. The universe of sets, of which the cumulative hierarchy is but a part, is, globally, inconsistent. Obviously, therefore, we need to use a paraconsistent logic when exploring it.

Let us call Frege's axioms above, based on a paraconsistent logic, *naive set theory*. The paraconsistent logic had better not contain the absorption principle, since otherwise trivialization due to Curry paradoxes will occur.[33] Thus, we may take this to be an absorption-free De Morgan logic. An important result due to the Australian logician Brady (1988), is that for many De Morgan logics without absorption, naive set theory is non-trivial.[34]

In naive set theory we can provide for all the set-theoretic constructions the working mathematician, including the category theorist, needs (unions, intersections, pairs, ordered pairs, functions, infinite sets, power sets, etc.), including those she cannot get in ZF (complementation, the universal set, etc.) (cf. Routley, 1977).[35] Whether these extra resources allow for any interesting and novel mathematical results, for example whether the category of categories (which, of course, has itself as a member) has any interesting category-theoretic properties, I do not know.[36]

9. Goedel's theorem

In this article I have concentrated mainly on technical issues. It is clear, however, that paraconsistent logic has important philosophical ramifications. By showing that it is possible to be inconsistent whilst remaining coherent, paraconsistent logics challenge any philosophical position which is based on an assumption of the necessity of consistency. Thus, for example, standard accounts of rationality, of existence, of motion, and so on, all depend on the unargued assumption that inconsistency is not to be countenanced. I shall not attempt to go into these issues here.[37] But I wish to mention briefly one example of the philosophical import of paraconsistent logic, since it is very closely connected with a number of the issues I have already discussed.

In the previous sections I have referred several times to the logical paradoxes. These, and the self-reflexivity inherent in them, are obviously closely connected with Goedel's theorem. Goedel's theorem has been thought to have many philosophical implications: the death of Hilbert's programme in the foundations of mathematics; the falsity of mechanism in the philosophy of mind; the inherent vagueness of the notion of natural number.[38] But Goedel's theorem about unprovability applies only to *consistent* theories.[39] Yet each of the philosophical applications promotes the unprovability into a universal result. And in each case, in the light of paraconsistent logic, we can see that inconsistency must be countenanced.[40] For example, if our conceptual apparatus of sets is inconsistent then so are both mind and mathematics. Paraconsistency therefore sweeps away, in a single blow, several philosophical arguments.

With this rather swift glance at the broader implications of paraconsistency, I must finish. I hope I have done enough to give you an idea of the subject and make its literature more accessible. I hope, however, that I have been able to do more than this, and that I have whetted your interest in it. Since paraconsistent logic is such a young subject, there is much that is unknown and requires research. And since the applications of the subject are so broad and deep, I think it likely that researchers in the subject will find further rich rewards there.

Notes

1. This talk was an invited address at the conference Interlogicon 87, an inter-disciplinary conference on mathematical logic and related subjects held at the University of Durban, July 1987. I am very grateful to the organizers of the conference, and particularly Chris Brink and the Department of Mathematics of the University of Cape Town for the invitation. I am also grateful to the South African HSRC, whose funding made this visit possible. I wish to make it clear, however, that the visit should in no way be interpreted as support for the racist policies of the South African Government. I wholeheartedly concur with the sentiments expressed by the South African Journal, *Philosophical Papers*, that Apartheid is an infringement of human, civil and academic rights, and an affront to human dignity.
2. More comprehensive and detailed surveys can be found in Priest and Routley (1983) and (1984). Many of the issues I mention are further discussed in Priest (1987). I will refer to these sources in what follows as *SPL, IPL* and *IC* respectively. The collections of essays to which the first two of these belong provide a more detailed picture of the discipline. A distinctly South American survey of the subject can be found in Arruda (1977).
3. *SPL*, sec 3.1. This may explain why Jaskowski preferred discursive implication to strict implication, which also satisfies *modus ponens*, but which he seems to ignore.
4. Many of these have been investigated, especially by Polish logicians and da Costa. See *IPL*, sec 3.
5. For example, that of the Canadian logicians Schotch and Jennings, *ibid.*
6. See da Costa (1974), *IPL* sec 3, and *SPL* secs 2.2 and 3.2.
7. Da Costa's original approach was proof-theoretic; the semantics came later. However, as throughout this article I shall concentrate on semantics rather that proof theory. Such an approach is, I think, much more illuminating.
8. Or, more generally, a set of formulas and a formula. For simplicity I will deal only with the single premise case.
9. A number of these are discussed in Epstein and Walton (1979). An elegant generalization of Smiley's filter is proposed by the erstwhile South African logician Tennant in (1984).

10. It is worth noting that although filter logics characteristically result in the failure of transitivity, filters can be found which respect it. See Dunn (1980).
11. A more usual, and equivalent, form of the disjunctive syllogism in the context of De Morgan logics is: $\{\alpha, \neg\alpha \vee \beta\} \vdash \beta$. (*Modus ponens* for 'material implication'.) It should be noted that the disjunctive syllogism also fails in positive-plus logics. However, I chose this as the mark of De Morgan logics since in these, unlike the positive-plus logics, it is the only major principle of inference concerning negation guaranteed to fail. (Indeed, it is equivalent to *ECQ*.) The philosophical debate concerning De Morgan logics has also promoted their counter-examples to it to centre stage.
12. There are, in fact, a number of different semantics for these logics including, notably, Routley-Meyer semantics. (See *IPL* sec 3, and *SPL* secs 2.3 and 3.3.) The algebraic semantics provide a convenient focus for the present occasion, however.
13. See Anderson and Belnap (1975), ch 3 sec 18. De Morgan lattices have also been used in other contexts, where they have been called distributive involution lattices or quasi-Boolean algebras. See Anderson and Belnap, *loc cit* p 194.
14. Again, for simplicity, I restrict myself to the single premise case.
15. For technical reasons, it is often convenient to add to the algebra a binary 'consistency' operator, o, called fusion. It is then $(aob^*)^*$ that represents $\rightarrow$ algebraically.
16. Notably by Routley, Meyer and their associates. See Routley *et al* (1982). Their investigations have, however, used mainly (im)possible-world semantics rather than algebraic semantics. For algebraic investigations, see e.g., Meyer and Routley (1972) and Priest (1980).
17. This is true of the logic RM. It is even clearer in the case of RM3, a logic which has played a fundamental role in the investigations of naive set theory. Paraconsistent Lukaciewicz many-valued logics, such as $Ł_{\aleph}$ with designated values [.5, 1], are also in this class, as is the logic of *IC* ch 6.
18. Lin (1986). For other algorithms, see Dunn (1976), Belnap (1977) and Bolan (1985).
19. See, e.g., Jeffrey (1981), ch. 2.
20. By members of the Automated Reasoning Project at the Australian National University. See Thistlewaite *et al* (1987).
21. It should be noted, however, that resolution can be interpreted as an application of transitivity $\{\alpha \rightarrow \beta, \beta \rightarrow \gamma\} \vdash \alpha \rightarrow \gamma$, for a genuine implication operator (not material implication), and therefore as quite correct. In this way, *Prolog*, say, can be seen as a perfectly correct fragment of a De Morgan logic. If resolution is interpreted in this way, of course, it is no longer true that every formula can be put in clausal form.
22. Notably, the Anderson Belnap systems E and R. See Urquhart (1984).
23. See Tarski (1956), sec. 1.
24. See, e.g., Perlis (1985).
25. See Priest (1984a) and *IC* ch 1. Another suggestion is to jettison the law of excluded middle. This is just as problematic. See the references just cited.
26. Whether truth is to be understood in its simple, Tarskian, sense, the constructive sense of the verificationists, or the truth-in-a-possible world form of Montague semantics.
27. See Priest and Crosthwaite (1988) and *IC* ch. 9.
28. Curry actually showed it for set theory. But the situation in semantics is similar. See n 33.
29. The Anderson Belnap logics E and R both contain absorption for example.
30. Of course, Frege's axioms appeared somewhat different, being in a second order theory in which ε was defined. Essentially, the claim is right enough, however.
31. But it is equally the case in formal semantics. See *IC* ch. 3.
32. See *IC* ch 2, and also Bell (1981).
33. Let c be $\{x; x \varepsilon x \rightarrow \beta\}$, for arbitrary β. The abstraction scheme then gives us $c \varepsilon c \leftrightarrow (c \varepsilon c \rightarrow \beta)$. The rest of the proof of β is as in the semantic case.
34. Trying to characterize the extent of inconsistency is, however, a different and unsolved matter.
35. Note, however, that some care has to be taken in the definition of some of these constructions if they are to be shown to have the usual properties; and some things will strike mathematicians, accustomed as they are to ZF, as rather strange. For example, it is possible to produce sets other than the empty set that have no members. It cannot be inferred from $\neg\exists x\varphi(x)$ and $\neg\exists x\psi(x)$ that $\forall x(\varphi(x) \leftrightarrow \psi(x))$.
36. It should also be noted that not all of Cantorian set theory appears to be forthcoming in the theory. For example, standard proofs of Cantor's theorem break down in De Morgan logics. Working mathematicians might see this as something of a blessing.
37. On the question of rationality, see Priest (1986) and *IC* ch. 7; on motion, see Priest (1985) and *IC* ch. 12; on being, see Routley (1980), esp. ch. 5.
38. The first is folklore. For the second, see Lucas (1961) and the literature it spawned [reviewed in Chihara (1972)]. For the third, see Dummett (1963).
39. In fact, there are even consistent theories, based on De Morgan logics, strong enough to represent all recursive functions which can prove their own absolute consistency, as Meyer (1975) has shown. Goedel's first undecidability theorem still, however, applies.
40. Indeed, it can be argued that versions of Goedel's theorem itself, show that certain situations must be inconsistent. See Priest (1979) sec II, and Priest (1984) secs 6, 7.

References

Anderson, A. & Belnap, N. 1975. *Entailment*. Princeton: Princeton U.P.

Arruda, A. 1977. A survey of paraconsistent logic, in Arruda *et al* (Eds.), *Non classical logic, model theory and computability*. North Holland.

Batens, D. 1980. Paraconsistent extensional propositional logics, *Logique et Analyse*, vol. 90–91, p. 195–234.

Bell, J. 1981. Category theory and the foundations of mathematics. *Br. J. Philos. Sc.*, vol. 32, p. 349–358.

Belnap, N. 1977. A useful four-valued logic, in J. Dunn and G. Epstein (Eds.), *Modern uses of multiple-valued logic*. Dordrecht: Reidel.

Bolan, A. 1985. *A relevant reasoner*. Research Report 18, Department of Philosophy, RSSS, Australian National University.

Brady, R. 1988. The non-triviality of dialectical set theory, in Priest *et al* (1988).

Chihara, C. 1972. On alleged refutations of mechanism using Goedel's incompleteness result. *J. Philos.*, vol. 69, p. 507–526.

Curry, H. 1942. The inconsistency of certain formal logics. *J. Sym. Logic*, vol. 7, p. 115–117.

Da Costa, N. 1974. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, vol. 15, p. 497–510.

Dummett, M. 1963. The philosophical significance of Goedel's theorem. *Ratio*, vol. 5, p. 140–155.

Dunn, J.M. 1976. Intuitive semantics for first degree entailment and 'Coupled trees'. *Philos. St.*, vol. 29, p. 149–168.

Dunn, J.M. 1980. A sieve for entailments. *J. Philos. Logic*, vol. 9, p. 41–57.

Epstein, R. & Walton, D. (Guest Eds.) 1979. *Philos. Stud.*, vol. 36, 2.

Jaskowski, S. 1969. Propositional calculus for contradictory systems. *Studia Logica*, vol. 24, p. 143–157.

Jeffrey, R. 1981. *Formal logic, its scope and its limits.* 2nd ed. McGraw Hill.

Lewis, C. & Langford, C. 1959. *Symbolic logic.* 2nd ed. Dover Publications.

Lin, F. 1986. Tableau systems for some paraconsistent logics. To appear.
Lucas, J. 1961. Minds, machines and Goedel. *Philos.*, vol. 36, p. 112 – 127.
Meyer, R. 1975. The consistency of arithmetic. Typescript. Abstracted as 'Relevant arithmetic', in the *Bulletin of the Section of Logic of the Polish Academy of Sciences*, vol. 5(4), 1976.
Meyer, R. & Routley, R. 1972. Algebraic analysis of entailment. *Logique et Analyse*, vol. 15, p. 407 – 428.
Perlis, D. 1985. Languages with self-reference I: Foundations. *Artificial Intelligence*, vol. 25, p. 301 – 322.
Priest, G. 1979. Logic of paradox. *J. Philos. Logic*, vol. 8, p. 219 – 241.
Priest, G. 1980. Sense, entailment and *modus ponens*. *J. Philos. Logic*, vol. 9, p. 415 – 435.
Priest, G. 1984. Logic of paradox revisited. *J. Philos. Logic*, vol. 13, p. 153 – 179.
Priest, G. 1984a. Semantic closure. *Studia Logica*, vol. 43, p. 117 – 129.
Priest, G. 1985. Inconsistencies in motion. *American Philos. Quar.*, vol. 22, p. 339 – 346.
Priest, G. 1986. Contradiction, belief and rationality. *Proceedings of the Aristotelian Society*, vol. 86, p. 99 – 116.
Priest, G. 1987. *In contradiction*. Nijhoff.
Priest, G. 198+. Reasoning about truth. *Artificial intelligence*, to appear.
Priest, G. & Crosthwaite, J. 1988. Relevance, truth and meaning, in R. Sylvan & J. Norman (Eds.), *Directions of relevant logic*. Nijhoff.
Priest, G. & Routley, R. 1983. Systems of paraconsistent logic, ch. 3 of *On paraconsistency*. Research Report 13, Philosophy Department, RSSS, Australian National University. Reprinted as ch. 5 of Priest *et al* (1988).
Priest, G. & Routley, R. 1984. Introduction: Paraconsistent logics. *Studia Logica*, vol. 43, p. 3 – 16.
Priest, G., Routley, R. & Norman, J. 1988. *Paraconsistent logic*. Philosophia Verlag.
Routley, R. 1977. Ultralogic as universal? *Relevance Logic Newsletter*, vol. 2, p. 50 – 90, and p. 138 – 175. Reprinted as an Appendix in Routley (1980).
Routley, R. 1980. *Exploring Meinong's jungle and beyond*. Department of Philosophy, RSSS, Australian National University.
Routley, R. *et al*, 1982. *Relevant logics and their rivals*. Ridgeview.
Smiley, T. 1959. Entailment and deducibility. *Proceedings of the Aristotelian Society*, vol. 59, p. 233 – 254.
Tarski, A. 1956. The concept of truth in formalised languages, ch. 8 of *Logic, semantics and meta-mathematics*. Oxford: Oxford UP.
Tennant, N. 1984. Perfect validity, entailment and paraconsistency. *Studia Logica*, vol. 43, p. 181 – 200.
Thistlewaite, P., McRobbie, R. & Meyer, R. 1987. *Automated theorem proving in non-classical logics*. Wiley.
Urquhart, A. 1984. The undecidability of entailment and relevant implication. *J. Sym. Logic*, vol. 49, p. 1059 – 1073.