IGOR URBAS

PARACONSISTENCY

Paraconsistent logics are distinguished by their suitability for reasoning from inconsistent premises without resultant collapse into triviality. They are therefore of immediate interest to those philosophers, among them some "dialectical" thinkers, whose views in some way commit them to accepting contradiction.

This is not to say, however, that such a strong philosophical position is required in order to prefer paraconsistent logics to their less tolerant competitors. On the contrary, the range of motivations for paraconsistency is as broad as the range of paraconsistent logics themselves. A rough classification of motivating positions is as follows.

The Dialethic Position. This is the strong position of those who believe that some contradictions (pairs of mutually inconsistent propositions) are true. Coupled with the usually held belief that not every proposition is true (non-triviality), this view rejects as not even truth-preserving the classically validated inference from inconsistent premises to arbitrary conclusions. A defense of this motivation for paraconsistency is to be found in Priest and Routley [8].

The Pragmatic Position. This position is not committed to the truth or otherwise of contradictions. Rather, it simply recognizes that many of our beliefs, judgements, scientific theories (including highly prized ones), legal codes and so on actually turn out to be inconsistent. Indeed, the likelihood of inconsistency seems to increase simply with the richness or expressiveness of theories, as in the case of semantic and other paradoxes engendered by self-referentiality. Thus, the pragmatic approach is to not abandon theories once they are discovered to be inconsistent, but to accomodate them (at least while there is no better alternative) by means of a logic which continues to function plausibly under the burden of inconsistency. An important proponent of this motivation for paraconsistency is the Brazilian logician da Costa (see [5]).

Studies in Soviet Thought **39**: 343—354, 1990.

Independent Positions. This last category includes any position which, for reasons not directly related to inconsistency, nonetheless endorses some kind of restriction on the consequences of contradictory premises. A good example is provided by relevant logicians, who insist on a connection of relevance or commonality of content between premises and conclusion in any valid inference. Proponents of relevantism not driven by questions of inconsistency include the pioneers Anderson and Belnap (see [1]).

Given this diversity of motivating positions, it is not surprising that paraconsistent logicians have generally rested content with the most minimal — and thus most widely acceptable — of formal definitions.

Preliminary Definitions D0. A *theory* T is a set of sentences — expressed in some (normally formal) language — which is closed under the consequence relation of the underlying logic L of T; i.e., if sentences $A_1, \ldots, A_n$ are in T, and B is a consequence of (is deducible from) $A_1, \ldots, A_n$ according to L, then B is also in T. A theory is *inconsistent* if it contains some sentence A together with its negation $\neg A$. (Note: logics usually incorporate a single primitive negation connective $\neg$; where several such connectives are present or definable, it is reasonable to demand that one of these be identified as the intended "contradictory-forming operator" in terms of which inconsistency is defined). Finally, a theory is *trivial* if it contains every sentence of its language; otherwise it is *non-trivial.*

Definition D1. A logic L is *paraconsistent* if it can support inconsistent non-trivial theories. Equivalently, L is *paraconsistent* if not every sentence B is a consequence according to L of inconsistent premises A and $\neg A$ (in symbols: $A, \neg A \nvdash B$). Otherwise, L is *explosive* (A, $\neg A \vdash B$).

The first thing to note about this definition of paraconsistency is that it is neither very restrictive nor substantive. It serves negatively to disqualify explosive logics as clearly inadequate for reasoning in inconsistent situations, but it does nothing positively to illuminate those that remain. All that it tells us about paraconsistent deducibility is that a certain inference is rejected, leaving it otherwise entirely open which inferences are to be accepted.

In response to this vacuum, positive and less formal constraints have periodically been appended to the definition. The Polish logician Jaśkowski, whose 1948 paper [6] is generally agreed to be the earliest formal exposition of "inconsistency-tolerating" logics, added that they should be rich enough to enable "practical inference" and should have an "intuitive justification." Over a decade later, da Costa adopted Jaśkowski's project and founded the South American school of paraconsistency — the name is credited to his Peruvian colleague Quesada — in the process adding several conditions of his own. The most important of these is that paraconsistent logics should approximate classical logic insofar as satisfaction of Definition D1 allows (see, for example [5]).

While some such additional conditions are perhaps desirable and even necessary in order to thin out the unmanageably large range of logics not excluded by D1, it is reasonable first to enquire whether this definition itself might be amenable to improvement. For clearly there are logics which satisfy the letter of D1 while brazenly flouting its spirit.

Example E1. Johansson's "minimal calculus" (see [7]) is obtained from positive intuitionistic logic by adding a falsity constant F and defining a "minimal" negation as follows: $\neg A =_{\mathrm{df}} A \supset F$. Although $A, \neg A \vdash B$ is avoided by Johansson's logic, it delivers $A, \neg A \vdash \neg B$.

Example E2. In [2], Arruda and da Costa introduced a family of five logics designed to cope with the paradoxes normally emergent in "naive" set theories. One of the stronger logics, J_5, turned out to be explosive in yielding $A, \neg A \vdash B$; of the remainder, J_2 to J_4 had the curious property that, while avoiding $A, \neg A \vdash B$, they delivered $A, \neg A \vdash B \supset C$ (see [3]).

Clearly, such logics are unsatisfactory for the purpose of supporting inconsistent theories, even though they technically satisfy Definition D1. Any inconsistent theory based on Johansson's logic contains every expressible sentence of the form $\neg B$. In a clear intuitive sense, such a theory is therefore "globally inconsistent" — every sentence in the theory is contradicted by its negation, also present. Similarly, inconsistent theories based on J_2 to J_4 include every expressible implication $B \supset C$. In the case of "naive" set theories based on these logics, this leads to the disastrous result that all sets are identical: Theorem 4 of [3].

These examples suggest that Definition D1 needs to be strengthened so as to exclude not only absolute explosiveness, but also explosiveness which is specific to particular operators or connectives. This can be achieved as follows.

Preliminary Definition D2. Let * represent a unary (respectively, binary, etc.) connective. Then a theory is **-trivial* if it contains all expressible sentences having * as principal connective. Otherwise it is **-non-trivial.*

Definition D3. Let * be as in Definition D2. Then a logic L is **-paraconsistent* if it can support inconsistent *-non-trivial theories. Equivalently, L is **-paraconsistent* if $A, \neg A \nvdash *B$ (respectively, $A, \neg A \nvdash B*C$, etc.) in L. Otherwise L is **-explosive.*

Definition D3 seems a good advance over its predecessor D1. By requiring that logics be not only paraconsistent but *-paraconsistent (for various connectives *), we ensure that they do a better job of containing the deductive damage caused by inconsistency. Accordingly, their inconsistent theories are more likely to prove satisfactory than the nearly trivial ones discussed in connection with Examples E1 and E2.

The pressing question that arises now is for which connectives * is *-paraconsistency a reasonable requirement.

Primitive Sentential Connectives. A modest beginning is to require *-paraconsistency for * ranging over the usual primitive sentential connectives $\neg$, $\supset$, & and $\vee$. The case for $\neg$ and $\supset$ is illustrated by Examples E1 and E2. Similar considerations support the extension to &-paraconsistency and $\vee$-paraconsistency: it is in fact possible to contrive logics which are paraconsistent but &-explosive or $\vee$-explosive.

Quantifiers, Modal and Other Primitive Connectives. The extension of *-paraconsistency to quantifiers so as to exclude the inferences $A, \neg A \vdash \forall xB$ and $A, \neg A \vdash \exists xB$ seems unobjectionable. However, the case of modal and other intensional operators is less clear. For example, an interesting philosophical case can be made for modal paraconsistent logics incorporating the inference $A, \neg A \vdash \Diamond B$, with the effect that in inconsistent (possible) worlds everything is possible (though perhaps not everything holds). Thus, in the absence of a uniform argument for extending *-paraconsistency to cover such connectives and operators, it seems preferable to concentrate on the more fundamental problem of ensuring a sufficiently strong form of paraconsistency at the sentential level.

Non-primitive (Definable) Sentential Connectives. Definable connectives present more of a problem. For example, *-paraconsistency for all of the primitive connectives $\neg$, $\supset$, & and $\vee$ is no guarantee of *-paraconsistency for familiar definable connectives like equivalence, $B{*}C = (B \supset C) \& (C \supset B)$, and "material implication," $B{*}C = \neg B \vee C$. Nor is, say, $\vee$-paraconsistency sufficient to secure *-paraconsistency for complex disjunctions such as $B{*}C = B \vee (B \vee C)$.

On the other hand, there are some definable connectives for which *-paraconsistency is not obviously desirable, such as the unary connectives ${*}B = B \supset B$ or ${*}B = \neg B \vee B$. The latter formulas are *theorems* (in symbols: $\vdash A$) of a great many logics, including established paraconsistent ones. Since most such logics also incorporate a rule of Weakening (allowing the move from $\vdash A$ to $\Gamma \vdash A$ for any sequence of formulas Γ), it follows that $A, \neg A \vdash B \supset B$ and $A, \neg A \vdash \neg B \vee B$ are in fact widely accepted inferences. Nor is any obvious harm occasioned by the presence of logical theorems in theories, including inconsistent ones. (Theories containing all the theorems of their underlying logics are called *regular*).

The problem, then, is to find a way of extending *-paraconsistency to cover some, but not all, definable connectives, excluding at least theorem-generating ones. One possibility is to exploit the differences in generality between formulas through which connectives are (syntactically) defined. One formula can be said to be less general than another of the same form if the first can be obtained from the second by uniform substitution of sentential variables but not conversely. Thus, $B \supset B$ is less general than $B \supset C$; $\neg B \vee B$ is less general than $\neg B \vee C$, and so on. The idea then would be to require *-paraconsistency only for connectives * defined by means of formulas in maximally general form.

Such a strategy would clearly exempt theorem-generating connectives, since formulas in maximally general form obviously contain no repetitions of sentential variables, whereas the theorems of virtually all logics do contain such repetitions (in the absence of devices like sentential constants). Unfortunately, however, too many other connectives would also be excluded, among them the earlier mentioned equivalence defined through $(B \supset C)$ & $(C \supset B)$, complex disjunctions such as that defined through $B \vee (B \vee C)$, and so on.

Thus, this strategy must be abandoned. The only workable solution which remains seems to be to exclude theorem-generating connectives from the demands of *-paraconsistency directly by stipulation.

Preliminary Definition D4. A connective * of a logic L is *theorem-generating* if all expressible sentences having * as principal connective are theorems of L.

Definition D5. A logic is *strictly paraconsistent* if it is *-paraconsistent for all (primitive and definable) sentential connectives *, excluding theorem-generating ones.

Definition D5 can be restated in several equivalent formulations.

Preliminary Definition D6. A theory is *strictly non-trivial* if it is *-non-trivial for all (primitive and definable) sentential connectives *, excluding theorem-generating ones.

Definition D7. A logic is *strictly paraconsistent* if it can support inconsistent strictly non-trivial theories.

Definition D8. Let B be any non-theorem of a logic L. Then L is *strictly paraconsistent* if there is some sentence B' obtained from B by uniform substitution such that $A, \neg A \nvdash B'$ in L.

Definition D5 and its equivalents are interesting in a number of respects. First among these is that, although strict paraconsistency is a substantially more restrictive concept than the mere paraconsistency of Definition D1, most logics which have been constructed specifically to be "inconsistency-tolerating" (such as the C-systems of da Costa [5]) in fact satisfy the more restrictive definition. So too, for that matter, do relevant logics (such as those of [1] and [9]). Indeed, the logics which are excluded by D5 are primarily:

(i) explosive logics, such as classical and intuitionistic logics, which were already excluded by D1;

(ii) nearly explosive logics, such as Johansson's, which were never intended to be paraconsistent anyway;

(iii) failed attempts like the $\supset$-explosive J_2 to J_4 of Arruda and da Costa [2]; and

(iv) contrived logics, which satisfy D1 but are &-explosive or $\vee$-explosive, or *-explosive for some definable (but not theorem-generating) connective *.

Thus, strict paraconsistency does not represent a departure from established paraconsistent tradition so much as a clearer and more informative formulation of its theoretical goal.

Secondly, Definition D8 shows particularly well how close paraconsistency, when taken seriously, comes to relevance. For it is not difficult to show that — except in the case of conclusions which are theorems — strictly paraconsistent inferences from contradictory premises satisfy the relevant requirement of *shared variables* (commonality of content) between premises and conclusions. The argument runs as follows. Let $A, \neg A \vdash B$ be a valid inference of some strictly paraconsistent logic L, where B is a particular formula which is not a theorem. Logics are standardly constructed so as to permit uniform substitution (either through an explicit substitution rule or by means of axiom and rule schemata which ensure the same effect). But if the premises and conclusion of the above inference share no sentential variable, then those variables occurring in B can be replaced with no effect on those occurring in A and $\neg A$, yielding the inference $A, \neg A \vdash B'$ for any substitution instance B' of B. But this contradicts the definition of strict paraconsistency in D8; thus, if L is indeed strictly paraconsistent, then the premisses and conclusion of the above inference must share a variable.

Of course, variable-sharing in inferences from contradictory premises to non-theorematic conclusions is no guarantee of variable-sharing in other inferences. But it is difficult to envisage a logic — much less a motivation for one — which would satisfy the variable-sharing requirement in only this subset of inferences. Indeed, inferences from contradictions and inferences to theorems are the two paradigm cases of failure of variable-sharing in traditional logics. In eschewing the first variety of failure, then, one would expect to be left at most with the second.

Accordingly, a greater degree of variable-sharing can be ensured by generalising on the preceding definitions. The idea is that arbitrary conclusions (or arbitrary conclusions of specific non-theorematic forms) should fail to be deducible not just from contradictory premises, but from premises in general. (Since inconsistency therefore plays no specific part in the following definitions, it is best not to designate these as forms of paraconsistency, but to return instead to the concept of explosiveness).

Definition D9. Let Γ be a finite sequence of formulae, and let * represent a unary (respectively, binary, etc.) non-theorem-generating connective. Then a logic L is **-non-explosive from* Γ if it can support *-non-trivial theories containing Γ. Equivalently, L is **-non-explosive from* Γ if $\Gamma \nvdash *B$ (respectively, $\Gamma \nvdash B*C$, etc.) in L. Otherwise L is **-explosive from* Γ.

Note: Definition D9 is essentially the same as a definition of "destructiveness" given in Batens [4].

Definition D10. Let Γ be as in D9. Then a logic L is *strictly non-explosive from* Γ if it is *-non-explosive from Γ for all (non-theorem-generating) connectives *. Equivalently, L is *strictly non-explosive from* Γ if it can support strictly non-trivial theories containing Γ.

The following definition is also equivalent.

Definition D11. Let Γ be as in D9, and let B be any non-theorem of a logic L. Then L is *strictly non-explosive from* Γ if there is some sentence B' obtained from B by uniform substitution such that $\Gamma \nvdash B'$.

Note: The restriction to finite Γ in Definition D11 (and corresopondingly in D9 and D10) serves to exclude the case in which all substitution-instances of B are already in Γ.

A final step of generalisation is available.

Definition D12. A logic is *strictly non-explosive* if it is strictly non-explosive from Γ for all (finite) sequences Γ.

Clearly, a logic which is strictly non-explosive (and which allows uniform substitution) satisfies the variable-sharing requirement in all inferences except those in which the conclusion is a theorem. Moreover, most established paraconsistent logics (such as da Costa's *C*-systems) are not only strictly paraconsistent but strictly non-explosive (and allow uniform substitution). Thus, the *only* obstacle left on the road from paraconsistency to relevance (or, at least, to unrestricted variable-sharing, which is a cornerstone of relevance) is the proviso exempting inferences to theorematic conclusions in Definitions D9 onwards.

Since this proviso is therefore so pivotal, it is worth examining the arguments for and against its inclusion. A first argument in favour is that a regular theory (one containing the theorems of its underlying logic) is in no way undesirable, as opposed to a trivial or nearly trivial one. But the easy reply to this argument is that requiring a logic to be able to support irregular theories — which would be the result of deleting the proviso — is not an endorsement of irregularity any more than requiring it to be able to support inconsistent non-trivial theories is an endorsement of inconsistency. What is at issue is not so much the properties that theories *should* have as the range of properties that they *can* have — because this range reflects the deductive properties of their underlying logics. So the dispute is really over the rule of Weakening, which forces theories to be regular.

The most general form of Weakening (of premises) is the rule from $\Gamma \vdash A$ to $\Delta, \Gamma \vdash A$. (In the case that A is a theorem, Γ is the empty sequence). The classical and perhaps most intuitive argument for Weakening is that, if A is deducible from a set of premises Γ, then surely A is deducible from an enlarged set of premises containing Γ. Thus, retention of this rule is in accordance with da Costa's requirement that paraconsistent logics contain as much of classical logic as the prior requirement of "inconsistency-toleration" allows. The relevant response to this argument for Weakening is that it strains the meaning of "from" — the deduction of A in the latter case is still *from* Γ, despite the addition of possibly irrelevant and therefore unusable premises Δ. Hereafter, the dispute tends to degenerate into a protracted battle over fundamental theoretical terms like "deducibility," "entailment," "consequence," and the relations between them.

However, there is a more direct argument against Weakening for paraconsistent logics. It is that the presence of this rule renders it impossible to add anything but the most impoverished imitations of negation without thereby producing some version of explosiveness. For example, Weakening plays an essential part in yielding the inference A, $\neg A \vdash \neg B$ in Johansson's logic, rendering even this system with its "minimal" negation $\neg$-explosive. A similar story holds for da Costa's C-systems, as is documented in [8] and [10].

Example E3. In [10], it is shown that the C-systems lack elementary negation properties, such as the property of intersubstitutivity of provable equivalents in negated contexts. For example, if formulas A and B are provably equivalent ($\vdash (A \supset B) \& (B \supset A)$, abbreviated $\vdash A \equiv B$), then so should their negations be ($\vdash \neg A \equiv \neg B$). This fails spectacularly in the C-systems: they affirm $\vdash A \& \neg A \equiv \neg A \& A$ but not $\vdash \neg(A \& \neg A) \equiv \neg(\neg A \& A)$, to take just one example. Moreover, the addition of Contraposition rules RC ($A \supset B \vdash \neg B \supset \neg A$), which is sufficient to restore the intersubstitutivity property, or the weaker EC ($A \equiv B \vdash \neg B \supset \neg A$), which is both sufficient and necessary, collapses all but the weakest system C_ω into explosive classical logic.

Example E4. Even though C_ω can bear the addition of rules RC and EC without becoming explosive, there is still some room for doubt about the paraconsistent merits of the resulting systems, respectively, RC_ω and EC_ω. In [10], it was intended that the premises of these rules be logical theorems, but a case can be made for closing not just logics but also their theories under these rules, resulting in a property of intersubstitutivity of theoretic equivalents. Thus, if $A \supset B$ or $A \equiv B$ is in a theory, then so would be $\neg B \supset \neg A$. The problem is that Weakening delivers a flood of implications and equivalences to which the Contraposition rules could be applied. In the form $A \vdash B \supset A$ (which is how it appears in C_ω, and which is easily derived from the form given earlier), it is clear that Weakening adds to a theory T containing A every implication $B \supset A$ for arbitrary B. Now apply RC to get $\neg A \supset \neg B$. If T also contains $\neg A$ (and is therefore inconsistent), then *modus ponens* (in the instance $\neg A, \neg A \supset \neg B \vdash \neg B$) yields $\neg B$, and T is $\neg$-trivial. Similarly, if A and B are both in T, then Weakening yields $B \equiv A$. Apply EC to get $\neg A \supset \neg B$. Again, if T also contains $\neg A$ (and is therefore inconsistent), *modus ponens* yields $\neg B$. This does not quite amount to $\neg$-triviality, since it was assumed that B is in T and therefore not arbitrary — it leads rather to a (still undesirable) property which we might call $\neg$-*saturatedness*: for all B in T, $\neg B$ is also in T.

Although there are delicate philosophical arguments involved in applying Contraposition rules like RC and EC to theories over and above logics, Example E4 shows how Weakening once again gets in the way. For the application of these rules involves taking seriously the implications and equivalences which are asserted. Where Weakening is present, all sentences in a theory are asserted to imply each other and to be equivalent, which it is difficult to take seriously at all.

Of course, it is a familiar part of the argument for relevant logics that "implications" which conform to Weakening are *ipso facto* not to be taken seriously. The formulation of Weakening employed in Example E4 is none other than the positive "paradox of material implication", its negative counterpart being $A \vdash \neg A \supset B$, which leads quickly to the paraconsistency-destroying $A, \neg A \vdash B$. Attempting to avoid the second paradox while retaining the first requires the abandonment also of symmetry-guaranteeing rules like Contraposition, as well as desirable properties like the intersubstitutivity of provable equivalents. The result is logics which are paraconsistent only at the expense of important systemic properties and well-behaved connectives, especially negation. In short, the cost of retaining Weakening — and thereby the proviso exempting inferences to theorematic conclusions from the scope of truly thorough definitions of strict paraconsistency and strict non-explosiveness — is not worth it. Relevance, with its emphasis on variable-sharing as an indicator of shared content, supplies the "intuitive justification" for paraconsistency which the pioneer Jaśkowski was right to demand.

REFERENCES

[1] Anderson, A. R. and Belnap, N. D. Jnr., *Entailment: The Logic of Relevance and Necessity, vol. 1*, Princeton University Press, 1975.

[2] Arruda, A. I. and da Costa, N. C. A., "Sur le schéma de la séparation," *Nagoya Mathematical Journal*, v. 38, 1970, pp. 71—84.
[3] Arruda, A. I. and Costa, N. C. A. "Le schéma de la séparation et les calculs J_n," *Mathematica Japonicae*, v. 19, 1974, pp. 183—186.
[4] Batens, D. "Paraconsistent extensional propositional logics," *Logique et Analyse*, v. 90—91, 1980, pp. 195—234.
[5] da Costa, N. C. A. "On the theory of inconsistent formal systems", *Notre Dame Journal of Formal Logic*, v. 15, 1974, pp. 497—510.
[6] Jaśkowski, S. "Un calcul des propositions pour les systemes deductifs contradictoires," *Studia Soc. Scien. Torunensis*, Sec. A, 1, 1948, pp. 55—57. English translation: "Propositional calculus for contradictory deductive systems," *Studia Logica*, v. 24, 1969, pp. 147—157.
[7] Johansson, I. "Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus," *Compositio Mathematica*, v. 4, 1936, pp. 119—136.
[8] Priest, G. and Routley, R. "On paraconsistency," *Research Series in Logic and Metaphysics*, 1983, Australian National University, Canberra.
[9] Routley, R., Meyer, R. K., Plumwood, V. and Brady, R. T., *Relevant Logics and their Rivals 1*, Ridgeview, 1982, Atascadero, Cal.
[10] Urbas, I. "Paraconsistency and the *C*-systems of da Costa," *Notre Dame Journal of Formal Logic*, v. 30, 1989, pp. 583—597.

Instituto de Filosofía
CSIC
Madrid
SPAIN