DIDERIK BATENS

AGAINST GLOBAL PARACONSISTENCY

1. THE PROBLEM

The need for and use of paraconsistent logics[1] has recently gained general recognition; the logical community has been freed from the dogma that all knowledge is necessarily consistent. However, it turns out that paraconsistent logicians disagree on a central issue. There are those who want to replace classical logic by some paraconsistent logic, usually a relevant one. They consider the requirement of or search for consistency a deep-rooted philosophical mistake. For them inconsistencies are inherent in all thinking, possibly with the exception of some very limited and special contexts. Common among them seems to be the belief that there is a (most) general form of thinking, sometimes identified with 'natural' thinking, which is strictly[2] paraconsistent. I shall use the term "global paraconsistency" to denote this position.

There are other logicians, who want to attribute to each logic a particular set of domains in which it is adequate. Even if, fighting the stubbornness of traditional logicians, they tend to stress the abundance of inconsistent domains, they refrain from attributing to paraconsistent logics an exclusive, global or superior status.

At first sight, the latter position might appear rather conservative. If the consistency dogma has finally been rejected, why not go all the way? Belonging myself to the latter group, I tend to look for a more friendly qualification of the position: not conservative but anti-dogmatic. Once a dogma is rejected, the worst alternative is to replace it by an opposite dogma. The rational reaction, I think, is to find a theory that incorporates both the new insights and matches the merits of the older view. If things get complex this way, that's an inconvenience to overcome. Also, global paraconsistency remains committed to the belief in a 'correct' logic; it need not be universally applicable, but is seen as the correct general logic. In my view, this is just another dogma.

The disagreement is an important one. It has implications for the

Studies in Soviet Thought **39**: 209–229, 1990.

status of logic and for its relations to language and to the world. It has implications also to the way in which logic should be taught. Also, the problem is complex and far-reaching, as will soon be clear. In order to avoid superficiality, and not to scare away the reader for whom my general philosophical position on the matter is unacceptable, I shall restrict myself to presenting a series of partly independent arguments. They are all related to the following two aspects of the problem. (i) Global paraconsistency is neither necessary nor desirable; nor does it actually permit an adequate description of consistent domains, e.g., the meta-theory of most logics, including paraconsistent logics. (ii) There is no need to view natural language (and natural thinking) as a single paraconsistent system.

A minor complication is that global paraconsistency has adherents among logicians who have *no* fundamental objections against the consistency requirement. Notorious examples are Alan Anderson and Nuel Belnap, for whom paraconsistency is a side-effect of the search for a natural (relevant) implication.[3]

On many occasions I shall direct my arguments against Graham Priest's *In Contradiction*.[4] This is indeed an excellent book in which an elaborate version of global paraconsistency is spelled out in a careful, cautious, and rigorous way. Although the concentration on Priest's book may undermine the generality of some of my arguments, it will keep me from attacking a straw man.

2. THE UNIVERSALITY OF CONSISTENT TWO-VALUED SEMANTICS

By a logic I shall understand a system that may be characterized by a set of inferences of the form $A_1, \ldots, A_n \vdash B$. This restriction does not require monotonicity and does not exclude that the derivability relation be relevant.[5] Roman Suszko[6] and Newton da Costa[7] have proved independently and along different lines that any logic (in the defined sense) has a two-valued semantics. The proof by da Costa involves a specific type of saturated sets (replacing maximal consistent sets in the completeness proof). Suszko's proof is presented as a rather simple consideration. For the purpose of my argument, such an outline is sufficient: simply turn any expression "$A_1, \ldots, A_n \vdash B$" into the semantic clause "$v(A_1) = 0$ or ... or $v(A_n) = 0$ or $v(B) = 1$". By stipulating that $v: F \rightarrow \{0, 1\}$ (where F is the set of all formulas) and by defining "validity" and "semantic consequence" in the usual way, we arrive at a two-valued semantics for the logic under consideration.

The resulting semantics may be inelegant, it may be not be very helpful for understanding the meaning of the logical terms, and some logical terms may not be truth-functions. In some cases some of these features may be overcome, but in more cases some cannot. All this, however, is not essential for the point made by Suszko and da Costa.

The truth and falsehood, represented by "1" and "0" respectively, are Aristotelian in the sense that they correspond to the truth and falsehood involved in the traditional criterion for the correctness of an argument: it is impossible that the premises be true and the conclusion false.[8]

Although no reader will have misunderstood the definition of a logic presented at the outset of this section, the discussion that is to follow requires that I make a specification. Using the term "set," I referred to a classical, non-fuzzy and consistent set theory (as anyone does who speaks unqualifiedly). Most systems commonly called "logics" are logics in the above sense. Of course one might devise a system characterized by degrees of derivability or a system in which the derivability relation behaves inconsistently. For the vast majority of logical systems, including most paraconsistent ones, this is not the case.[9]

Given this interpretation of the term "set," it is trivial that any logic may be discribed in a classical metalanguage (that excludes inconsistencies). Even if the logic is paraconsistent, even if it is inconsistent in having contradictory theorems, there is no need for the metalinguistic negation to be paraconsistent. "$A_1, \ldots, A_n \vdash B$" will never obtain together with "$A_1, \ldots, A_n \nvdash B$." Also, $v(A) = 1$ and $v(A) = 0$ exclude each other: the two-valued valuations are functions in the classical sense and $1 \neq 0$. If it would turn our that $1 = 0$, we shall be in big trouble, by which I mean that we shall not only have to revise the semantics under consideration, but the very foundations of the underlying mathematics, a point to which I shall return.

The present result may be extended to any theory. Even if a theory is inconsistent, it may be described in a consistent metalanguage. The only requirement is that our definitions of such metalinguistic notions as "to be a theorem of the theory" or "to be derivable from a set of premises according to theory" warrant their consistency.

Let me add a clarification that concerns claims made in the present section, but should be remembered for the subsequent ones as well. I argued that any logic and theory (as qualified above) may be described adequately in a classical metalanguage. I do not exclude that such a language also contains a paraconsistent negation. Yet, for the present purposes there is no need for this, and clearly the classical negation should be used to define such symbols as "$\neq$" or "$\nvdash$". Also, I do not exclude that a classical metalanguage contains a relevant implication. Briefly, I call "classical" any language, the underlying logic of which contains a negation, say "$\neg$," such that A and $\neg$ A cannot be true together.

3. CLASSICAL EQUIVALENTS FOR PARACONSISTENT DESCRIPTIONS

Let me start this section by an example. Graham Priest's *In Contradiction* contains a logical system that is itself described paraconsistently. The basic idea is that the valuations do not assign a member but a subset of $\{0, 1\}$ to sentences. Priest excludes the empty set as a possible value for sentences. So, the possible values are $\{0\}$, $\{1\}$, and $\{0, 1\}$ and they may be interpreted respectively as truth, falsehood and joint truth and falsehood. Under this interpretation the semantic description is inconsistent: "false" meaning the same as "not true," some sentences being both true and not true. Negation is defined (in the symbolism I use in this paper) by

$$1 \in v(\sim A) \text{ iff } 0 \in v(A)$$
$$0 \in v(\sim A) \text{ iff } 1 \in v(A)$$

and disjunction by

$$1 \in v(A \vee B) \text{ iff } 1 \in v(A) \text{ or } 1 \in v(B)$$
$$0 \in v(A \vee B) \text{ iff } 0 \in v(A) \text{ and } 0 \in v(B)$$

The system also contains a truth-predicate, symbolized by "T" which at the propositional level may be characterized by

$1 \in v(TA)$ iff $1 \in v(A)$
if $0 \in v(TA)$, then $0 \in v(A)$

Actually, Priest analyses "T" as a predicate, but the two clauses above come out true, and the semantics of predicates is immaterial to my present point. Let me add, as a side-remark, that even these few data about the system show its coherence and its intuitive appeal.

It is trivial that the semantics may be transcribed as a three-valued one, with 1, 2 and 3 replacing $\{1\}$, $\{0, 1\}$ and $\{0\}$ respectively. Next, applying an algorithm I described elsewhere,[10] this three-valued semantics may be interpreted as a 'couple-semantics' with values $\langle 1, 0\rangle$, $\langle 1, 1\rangle$ and $\langle 0, 1\rangle$, in which 1 and 0 stand for two-valued truth and falsehood respectively, and in which a couple of values consists of the *two-valued* value of the sentence and the two-valued value of its negation. I summarize all this in the following schema:

Priest	$v(A)=\{1\}$	$v(A)=\{0,1\}$	$v(A)=\{0\}$
3-valued	$v(A)=1$	$v(A)=2$	$v(A)=3$
couple	$v(A)=\langle 1,0\rangle$	$v(A)=\langle 1,1\rangle$	$v(A)=\langle 0,1\rangle$
2-valued	$v(A)=1$ and $v({\sim}A)=0$	$v(A)=1$ and $v({\sim}A)=1$	$v(A)=0$ and $v({\sim}A)=1$

After discovering the schema, my reaction was: "Aha! When Graham says that a sentence is true only, viz. receives $\{1\}$, he means (in my terminology) that it is true, but that its negation is false. When Graham says that a sentence is both true and false, he means that both the sentence and its negation are true. Finally, when Graham says that a sentence is false only, he means that it is false, but that its negation is true." But actually the matter is slightly more complicated. There are indeed two differences between Priest's position and mine: one is expressed by the schema, the other concerns the interpretation of the statements in the schema.

The first difference is that Priest's formulation contains inconsistencies at the semantic level, whereas mine does not. Any true (in both senses) inconsistent statement A will be described as being both true and false, hence both true and not true, by him, whereas it will be described consistently by me, viz. by saying that both the statement and its negation are true (whereas neither is at the same time not true).

The second difference is a bit more complex. Let me spell out the problem by considering the four following statements:

(1) $v(A) = \{1\}$
(2) not $v(A) = \{0\}$
(3) $v(A) = 1$ and $v(\sim A) = 0$
(4) not $v(A) = 0$

(1) and (3) are equivalent formulations, and so are (2) and (4). Moreover, the odd-numbered statements entail the even-numbered ones. If (1) and (2) are part of Priest's (paraconsistent) semantics, then the "not" in (2) is paraconsistent. But as (2) and (4) are equivalent formulations, so is the "not" in (4). In other words, by presenting the equivalence of the four formulations in this way, I did not show that Priest's paraconsistent formulation is equivalent to my classical formulation. And, of course, it is impossible to show this: the classical formulation is stronger.

However, given that the (paraconsistent) two-valued formulation does not contain any inconsistencies, we may replace it by a classical formulation that uses the same terms but has a different underlying logic. This classical formulation is obviously *stronger* than the paraconsistent one, but it is easily seen that it is a correct semantic description of the logical system.

In a sense, the classical two-valued semantics is simply a reinterpretation of the paraconsistent two-valued semantics mentioned in the schema: basically, metalinguistic disjunction and conjunction are kept identical and metalinguistic negation is interpreted classically. But given this, what about the other formulations mentioned in the schema? In its classical interpretation the two-valued formulation is still equivalent to the other three formulations. In other words, we can go all the way back to Priest's semantics. There we had inconsistencies, e.g., an inconsistent sentence will be both true and not true, and these need of course special care. A simple solution is to keep the original paraconsistent negation only where it is required, viz., in metalinguistic inconsistencies, and to write it in a special way, say in capitals. Once this is done, such expressions as "true and NOT true" (i.e., "true and false") behave consistently themselves. Needless to say, there is no problem at all at the level of the values $\{1\}$, $\{0, 1\}$ and $\{0\}$.

My reasoning concerning Priest's logic may be generalized as follows. There is an algorithm that enables us to turn (certain) paraconsistent meta-logical systems into consistent two-valued (as well as consistent many-valued) meta-logical systems. The latter systems are equivalent to the former both under a classical and under a paraconsistent interpretation, but the two series of formulations are not identical in that the classical one is stronger. I guess that there are paraconsistent meta-logical systems to which the algorithm cannot be applied, e.g., because "validity" or "derivability" behaves inconsistently in them, but I cannot give an example. The algorithm does the job for all paraconsistent systems I looked into.

I already remarked that the classical descriptions are stronger than the paraconsistent ones. In general, this feature may be expected to be an advantage: talking classically, I am able to say all Priest is able to say, but *not conversely*.[11] I shall spell out this advantage in subsequent sections. For the moment, let me point to the intuitive appeal of the classical description. Even if some domain is inconsistent, and hence so is its correct description, there is no reason why our reasoning about this description should be inconsistent as well. If it turns out, to take a simplistic example, that it is both raining and not raining, then clearly the world is inconsistent (and the "not" should be paraconsistent). But my viewpoint will enable me to say that I consider both sentences to be true (my description of the world is inconsistent) without having to say in the same breath that I also consider them to be false (my reasoning about the description is consistent).

4. PARACONSISTENT TRUTH AND FALSEHOOD

Let us again consider Priest's predicates "true" and "false," symbolized in his system by "T" and "F" respectively. The system should contain such predicates, as it is meant to be a logic for natural language, a language that is capable of describing itself, including its well-known paradoxes, in a non-trivial way. It should be stressed that the result is most interesting and impressive. Still, it turns out to have a weakness at a somewhat deeper level. In order to facilitate the discussion, *I shall, throughout the present section, use "true" and "false" in Priest's sense.*

Let us at once compare two sentences: an inconsistent sentence A that is both true and false (some inconsistent sentences are only false) and a true (consistent) sentence B. If we enumerate all true simple statements about the truth and falsehood of A, we arrive at the following list:

(1) $TA, FA, \sim TA, \sim FA, T\sim A, F\sim A, \sim T\sim A, \sim F\sim A$

For B we obtain

(2) $TB, \sim FB, F\sim B, \sim T\sim B$

Both lists contain redundancies, e.g., TA and $\sim FA$ are logically equivalent.

At first sight, the situation is quite satisfying: the lists seem to distinguish neatly between A and B, e.g. in that "FA" is present in the first list, whereas "FB" is absent from the second. And yet there is a problem. Talking *consistently* about the lists, we conclude that A is both true and false, whereas B is true but *not* false. I shall now show that we cannot express, within Priest's system, what *we* mean when we say that A is both true and false whereas B is true but not false.

That B is *not* false (in the classical sense of "not") cannot be expressed by "$\sim FB$." Indeed this sentence, which is already present in list (2), merely expresses that B is true, but not that it is *not* false. By all means, notice that "$\sim FA$" occurs in list (1), whereas A *is* false. Analogously for "$\sim T\sim B$" which, being equivalent to "$F\sim B$", fails to express the absence of "FB" (or of "$T\sim B$") from list (2). The same obtains for "$\sim\sim TB$", etc. One might try to get around the problem by such statements as $\sim(TB \,\&\, FB)$ or $\sim(TB \,\&\, \sim TB)$. But these happen to be theorems of Priest's logic. They are true, but so are $\sim(TA \,\&\, \sim TA)$ and $\sim(TA \,\&\, FA)$. There simply is no way to express, within this logic, that B is *not* false or that B behaves consistently, i.e. is *not* both true and false (twice in the classical sense of not). The reason why

I add "in the classical sense" is that Priest will read ~FA as "A is not false." But here the "not" is paraconsistent and this "not false" does not exclude "false."

In summary: it is possible to express by paraconsistent means that a sentence is both true and false. If it is only true, we may obviously refrain from stating that it is false, but we cannot state that it is *not* false (in the strong sense). This certainly *is* troublesome. Natural language clearly enables us to express the difference between the A and B from the example. Graham Priest's logic is intended to capture natural thinking, but fails to do so.

The reason for the trouble lies with paraconsistent negation. A and ~A may be true together, and that is exactly what we need if talking about an inconsistent domain. By the same token, paraconsistent (weaker than classical) negation cannot enable us to express that something, e.g., FB, is false in the classical sense. Even if, talking paraconsistently, we say that "FB" does 'not' occur in list (2), this statement does not *exclude* that "FB" occurs in the list; and if we say that list (2) is a complete enumeration of all sentences of a certain form, this statement does not exclude that (2) is also in this respect incomplete.

I have shown two things in this section: (i) Priest is not able to express that B is *not* false if "not" is meant classically, and (ii) Priest is not able to *express* the difference between A and B within his system. Other paraconsistent systems may not have ~(TA & FA) as a theorem; in some of these, ~(TB & FB) will obtain, whereas ~(TA & FA) will not. Hence, the analogue of objection (ii) will not apply to these systems. But the analogue of objection (i) does.

5. PARACONSISTENT MATHEMATICS?

I fear that many readers will feel that either I am making a mistake or else something really unusual is going on. After all, the difference between the A and B from the example is quite obvious with Priest's semantics.[12]

(1) $v(TA) = \{0, 1\}$
(2) $v(TB) = \{1\}$

Or is it not? The answer depends on whether the mathematics used in the semantics is classical or paraconsistent. Whether Priest is able to produce a paraconsistent mathematics at this moment is not important; one cannot solve all problems in the same book, and I certainly would not object to a (somewhat controlled) intuitive use of a paraconsistent mathematics.

If the mathematics is classical, the distinction between A and B is indeed clear-cut: v is a function and $\{1\} \neq \{0, 1\}$. But in this case I have a central philosophical objection. Priest sees natural language as a single system, capable of coherent self-reference and claims to articulate the logic of this system. He took all the trouble to show that his system may be used as underlying logic for its own description. This actually is the capital argument for his view on natural language. So, using classical mathematics would be cheating, for it would mean that (at an essential point) a system is smuggled in that cannot be incorporated in natural language as described by Priest. In other words, it would mean hiding that natural language, as conceived, is not capable of adequately describing itself.

So, suppose Priest is using a paraconsistent mathematics that is compatible with his own system and philosophy (hence does not contain classical negation). I will show that, in this case, the semantic system does not adequately express the difference between A and B (from the ongoing example). Needless to say, each of the following will be theorems in paraconsistent mathematics:

(3) $\{0, 1\} \neq \{1\}$
(4) if $v(C) = x$ and $v(C) = y$ then $x = y$

From (1)—(4) we are able to derive:

(5) $v(TA) \neq \{1\}$
(6) $v(TB) \neq \{0, 1\}$
(7) $v(TA) \neq v(TB)$

Alas, as the negation used to define "$\neq$" is paraconsistent, (7), though clearly true, does *not rule out* (8):

(8) $v(TA) = v(TB)$

This is exactly the point of paraconsistent negation: to allow for the possible truth of inconsistencies. Of course, if (8) were to obtain, so would (9):

(9) $\{1\} = \{0, 1\}$

This would mess up paraconsistent set-theory to such an extent that most paraconsistent logicians would find it uninteresting. But this is immaterial to the point I am trying to make, viz., that we are unable to rule out (8) or (9) with the help of paraconsistent negation. In other words, if the underlying mathematics is paraconsistent (1) and (2) do not adequately express the distinction between A and B.

I think that the present section is quite clarifying. I guess that the weakness of Priest's system is not apparent at first sight because we are used to identifying mathematical terms with those of classical mathematics. This holds all the more for terms like "identity" or "function," the definition of which seems far from anything like a negation.

I hasten to add that paraconsistent mathematics *is* an important issue and tentatives to develop it should be applauded. There are cases in which we need paraconsistent mathematics. But at the same time we need to recognize at which points it is less powerful than classical mathematics, and to which cases we should not apply it.

6. INCOMPATIBILITY AND EXCLUSION

Allow me to mention a last argument that refers directly to Graham Priest's book. The passage on which I shall rely is not extremely central, but I think the line of reasoning is not infrequent among adherents of global paraconsistency. This is the reason why I shall spell out the argument.

It is well-known that disjunctive syllogism is not a correct argument according to any paraconsistent logic:

$$\frac{\begin{array}{c} A \lor B \\ \sim A \end{array}}{B}$$

My understanding of this fact is very simple. In paraconsistent logics A and $\sim$A may be true together. If they are, the second premise is true, because it is $\sim A$ itself, and the first premise is true, because A is, and this irrespective of the truth or falsehood of B. So, clearly the disjunctive syllogism is not correct within paraconsistent logics.

Adherents of global paraconsistency need to argue that the disjunctive syllogism is incorrect in general (is at best correct in some very specific cases). But then why is it that for so many centuries so many people thought it correct? Priest argues that the disjunctive syllogism is never correct, but is confused with the following:

(1) If $A \vee B$ is rationally acceptable and A is rationally rejectable, then B is rationally acceptable.

Priest claims[13] that acceptance and rejection (or refusal to accept) are "incompatible" (p. 128), that "joint rational acceptance and rejection are not possible" (p. 142). I think it is not coherent for him to do so.

Priest does his utmost to fight a negation, say "$\neg$" that is strong enough to make $\neg A$ and A incompatible, and then introduces incompatibility at the 'higher level' of acceptance. Of course, "incompatibility" may be understood in several ways, but it is quite clear that it should be understood in a very strong sense in order for his argument to go through. It is equally obvious that the "not" in the quotation from p. 142 should not be a paraconsistent negation, but a classical one.

The extend of the trouble is best illustrated by the following consideration. As "incompatible" belongs to the natural language for which Priest is devising a logic, one may use the term to define a negation "$\neg$" as follows:

(i) $\neg A \vdash \sim A$
(ii) $\neg A$ and A are incompatible

I claim that "$\neg$," thus defined, is a classical negation. If it were a paraconsistent negation, then the meaning of "incompatible" would be too weak for Priest's argument. The least a sufficiently strong sense of "incompatibility" should warrant is that incompatibles cannot possibly be true together. But if "$\neg$" is not paraconsistent, it is classical in the previously defined sense.

It is easy to show that Priest's "incompatibility" is tied to the ex falso quodlibet. Suppose indeed that A were both rationally acceptable and rationally rejectable. Then $A \vee B$ would be rationally acceptable and hence, by (1), so would be B. In other words: if A were both rationally acceptable and rejectable, everything would be rationally acceptable. (Presumably everything would also be rationally rejectable in view of the following facts: (i) the rejection of A entails the rejection of A & B, and (ii) the acceptance of A and rejection of A & B entails the rejection of B). As the previous result is a consequence of the incompatibility of acceptance and rejection, I cannot see a general meaning for "incompatible" that would not connect incompatibles with triviality.

7. EMPIRICAL FALSEHOOD

If a logic is to be applied in methodological contexts, the very least one should expect is that it allows for the possibility that certain observational statements form a problem for theories. There has been, and unfortunately still is, a multiplicity of views on the relation between theory and observation. Presumably the simplest such view is to be found with logical empiricism and with Popper. This view simply presupposes classical negation and rules out paraconsistency. But even views that leave room for the latter require classical negation at some point.

A first possibility is that the theory be formulated in a strictly paraconsistent language, but that the 'observation language' contain a classical negation that enables us to express the observed absence of some fact. A second possibility is that no classical negation be present, but that we may consider some observation as sufficient to *reject* some prediction. Here the reasoning is similar to the one in section 6: "rejection" enables us to define classical negation. A still further possibility, partly overlapping with the former, is that inconsistencies appear at the level of our observations.[14] In some such cases it will be desirable that the theory itself enable us to derive inconsistent predictions. As the negation employed to describe observations is paraconsistent, the observational report $\sim A$ will not in the presence of the prediction A constitute a problem for the theory. Nevertheless, we might formulate the problem as a meta-consideration: the theory predicts A, or perhaps even both A and $\sim A$, whereas A fails to appear among our observations. Given certain beliefs about the observational means and the meanings of the terms, this will indeed constitute a problem. However, here again we implicitly recur to a strong negation: the information we need is that A does not obtain in the strong sense of "not."

Let me consider a final possibility. More recent methodological views[15] stress the importance of empirical (and other) problems rather than restricting their attention to falsifications. The failure to predict or explain some facts taken to belong to the domain will be considered just as problematic as making incorrect predictions. With respect to such an approach, we need to express at the very least that some or other statement is not derivable from the 'union' of the theory and of a set of selected factual statements. So, in order to locate the empirical problems of the theory, we need to describe it by classical means (the arguments are of course the same as the ones used in previous sections).

I do not claim to have exhausted the possibilities, or to have fully discussed any one of them. Yet, I think the previous arguments show clearly enough where the problem with global paraconsistency lies.

8. CLASSICAL NEGATION

According to the classical analysis, the negation of a sentence is true just in case the sentence is false. The terms "true" and "false" have a specific meaning within classical semantics. I do not mean that we usually assign some interpretation to them, but that this meaning appears from the structure of the semantics itself, viz. from the fact that the semantic consequence relation is defined in terms of "truth in a model." Classical logic seems to offer an effective analysis of the meaning of classical negation: among other things, A and $\neg A$ exclude each other. But how are we to express this exclusion at the syntactic level? My claim is that the ex falso quodlibet does the job, and is the only effective means we have.

To adopt the ex falso quodlibet has dramatic consequences. Someone who asserts $\neg A$ is truly committed to the rejection of A: asserting A as well would commit one to triviality. The dramatic character lies in the fact that triviality constitutes the end of all thinking: if all sentences are (asserted to be) true, then no sentence carries any information,[16] no reasoning or thinking might add anything to our knowledge or improve it in any way. In view of this, classical negation enables us to express correctly that we reject some sentence: we can no longer accept it (except after having changed our minds).

Paraconsistent negation should not and cannot express rejection. It does not rule out the sentence that is negated and is intended *not* to rule this out. This is not an objection against paraconsistent negation, just as it is no objection to a violin that it is useless to hammer nails in the wall. But if we want to express the rejection of some sentence, we cannot recur to paraconsistent negation.

Adherents of global paraconsistency may introduce all kinds of terms at the syntactic level to specify the meaning of "to reject," and they may recur to all kinds of semantic machinery. But as long as the meaning of "to reject" is not tied up to triviality, it will necessarily be weaker than the classical meaning.[17] (Analogously for "incompatible," etc.) There is no way to get around this: the safety (against triviality) offered by paraconsistent negation is matched by the inexpressibility of classical rejection.

This then is the advantage of classical negation over paraconsistent negation. We do not always employ classical negation. Sometimes we can't. But if we do, we say more than if we don't. It's that simple.

9. INCOMPLETE DESCRIPTIONS

My argument from section 4 may now be generalized: if it is possible to describe some domain, e.g., some logical system, by classical means, this description will be more complete than the corresponding description by paraconsistent means. To be more explicit, in such cases the paraconsistent description is unacceptably incomplete.

Consider statement (1), which expresses that "$\supset$" is not an 'implication' in Anderson and Belnap's well-known system E.[18]

(1) B is not derivable from A and A $\supset$ B according to E.

Even if the "not" is paraconsistent, (1) is an informative statement: there is a proof of it which is by no means trivial. But (1) is also (provably) true if "not" is interpreted classically. In this reading, (1) *rules out* that detachment holds for material implication in E, and that is a more informative statement. *Every* logician will stop a student who is looking for a proof of the correctness of material detachment in E, and this shows that he or she means (1) to be true with "not" as classical. This is by no means less certain than most mathematical statements. If, apart from the proof of (1), there is a proof of its opposite, then something is damn wrong with the basics of present-day mathematics. Moreover, there is no reason to believe that the classical interpretation of the proof of (1) would be less reliable than any intended paraconsistent interpretation of it. Hence, the description of E by paraconsistent means, e.g., in a language which has E as its underlying logic, is unacceptably incomplete.

10. PARACONSISTENT DESCRIPTIONS OF CLASSICAL LOGIC

I have long thought I had to qualify certain claims made in this paper in view of my conviction that classical negations may be defined within a paraconsistent metalanguage. Nothing prevents us from writing (1) in such a metalanguage:

(1) $A, \neg A \vdash B$

I shall now show that (1) by itself does not warrant the ex falso quodlibet in its *classical* meaning.

By the same reasoning I used elsewhere, if (1) is stated in a paraconsistent metalanguage, it does not exclude (2).

(2) There are sentences C and D such that $C, \neg C \nvdash D$.

Although (1) certainly turns all $\neg$-inconsistent theories into trivial theories, it does not guarantee that "trivial" obtains its full sense. For one thing, all trivial theories are *identical* on the classical reading. Not so in the present case. In view of (2), trivial theories are only identical in the 'positive' sense (they share their theorems — all sentences) but may be different in the 'negative' sense (they need not share their non-theorems).

Apart from this systematic difference, there is also a more pragmatic one. Classical triviality is the end of all thinking: there is literally nothing to be said about a trivial theory, except that just everything is true according to it. But if (1) is formulated within a paraconsistent metalanguage, then clearly a lot may still be said. To take one example, one might define the proper theorems of a theory as the sentences that are theorems and are not non-theorems. The set of the proper theorems of a trivial theory is not in general trivial. It might then be used to characterize trivial theories; it might be used as the basis for defining the amount of information provided by sentences in the domain. More generally, one might map most statements about sets of theorems onto statements about proper sets of theorems.

The conclusion of all this is that (strictly) paraconsistent metalanguages do not enable one fully to describe classical logic. I hope the reader will appreciate the complexity of the situation. E.g., whenever I wrote "trivial" before, I should have added "in the classical sense" in order to avoid all misunderstanding. (I trust, though, that no reader misunderstood me.)

11. WHEN INCONSISTENCY CROPS UP

I have been arguing that paraconsistent systems are not adequately described by paraconsistent means. But suppose we describe them by classical means, thus excluding inconsistencies from the description, and that an inconsistency turns up. For example, we might find out that some formula both is and is not a theorem of some logic. No doubt we would be in trouble: the description collapses to triviality. So, is it not desirable after all to opt at once for a paraconsistent description?

For one thing, it certainly is *safer* to opt for a paraconsistent description. But is it wise to play things safe? Some logicians think so,[19] but they are mistaken. First of all, too much safety is a hindrance for progress. Popper has argued at length for bold conjectures in the 'empirical' sciences, but the situation for logic and mathematics is not different. Just imagine what mathematics would look like today if past mathematicians had opted for safety in view of the fact that we cannot be absolutely certain that 1 is different from 0, or that something might be wrong starting with some number with the statement that any natural number has a successor.

The safety provided by global paraconsistency is specifically undesirable. Under a classical description we are able to state that some domain (including logic and mathematics) is consistent. Under a paraconsistent description we are of course able to use the same word, or the same formulas, such as $\sim(A \ \& \ \sim A)$. But both the word and the formulas have different, weaker meanings. "Consistency" does not exclude inconsistency, as I argued before, and ~(A & ~A) hardly means anything at all in most paraconsistent systems (indeed it does not provide particularly salient information about any sentence in systems in which it is a logical theorem). In section 10, I showed that even classical triviality is beyond the reach of (strictly) paraconsistent viewpoints. To my mind, this forces us to give up all hope for ever expressing plain and simple old-fashioned consistency within global paraconsistency. The price for a global insurance against triviality is simply too high.

A final argument against safety is that, if things happen to go wrong, we can change our minds. If we try to describe some domain by classical means and it turns out to be inconsistent, why not simply correct our mistake and move to paraconsistent means? But maybe I am proceeding a bit too fast here. Advocates of global paraconsistency might indeed reason as follows. If you decide to move from some classical description to a paraconsistent one, you need the means to carry out this mental process; you need a *locus* — simplistically, some part of your brain — where you are able to coherently think your way through this process. But this process too might happen to involve inconsistencies; if you start off classically here, you may need yet another *locus*, and so on indefinitely.

It seems to me that the troubles to which such hierarchies lead are among the chief reasons why some people accept global paraconsistency. It is typical that *In Contradiction* contains a long argument against two such hierarchies. Priest convincingly rejects views according to which natural language is a hierarchy of separate systems or is a single system the interpretations of which form a hierarchy (pp. 23—28). My only disagreement concerns the fact that he seems to consider dialetheism, his particular brand of global paraconsistency, as the only alternative. I think there is another alternative for which there are good independent arguments, viz., 'contextualism'. The idea is that we do not depend on a fixed global system, which should be justified once and for all, but that we set up a specific context (involving meanings, relevant data, methodological instructions, etc.) whenever we meet a problem. The articulation and defence of this view falls beyond the scope of the present paper.[20]

In the present section, I argued that the risk of stumbling upon a metalinguistic inconsistency is worth being taken, and that safety is too expensive. I would like to add that even this particular safety does not pay. Suppose indeed that we find inconsistencies about theoremhood, derivability, semantic truth, or well-formedness. The effect would by no means be restricted to the meta-theory of some specific logical system; our deepest meta-logical means would be shattered. (If you doubt this, review sections 1—3 to see where the trouble might arise.) This would have two effects. On the one hand, the basic mathematical means (mappings, functions, etc.) would be ruined to such an extent that it would at once become impossible to establish the inconsistency we supposedly discovered. Moreover, the crisis would also affect paraconsistent mathematics; as I said before, nobody would stay content with a paraconsistent mathematics according to which both $v(A) = \{1\}$ and $v(A) = \{0, 1\}$.

12. FINAL COMMENTS

My central arguments purported to show that classical descriptions of many logical systems, including paraconsistent ones, are possible, and that, whenever this is so, paraconsistent descriptions are too poor to be adequate. The same applies to other domains: whenever a domain is consistent, a paraconsistent description is incomplete.

The central trouble is with the weakness of paraconsistent negation: inconsistencies do not constitute a *problem*, at least not a 'serious' one. The metalinguistic description of classical negation is beyond the reach of paraconsistency, and so is classical triviality. The fact that inconsistencies are unproblematic from a paraconsistent viewpoint has implications for the relation between dialectics and paraconsistency.[21]

Before ending this paper, I should repeat that fundamental disagreements usually cannot be solved in view of a few arguments. A basic disagreement concerns the possibility of a single, ultimate, overall and absolutely justified (paraconsistent) system. The contextual point of view is not affected by the shortcomings of global paraconsistency, is attractive for independent reasons, and does not require that any domain be in principle safeguarded against inconsistency. If this alternative is viable, as I believe, then we are able to stick out our necks and use the stronger, classical means to describe the domains that we believe to be consistent; if the belief turns out to be mistaken, we still may change our minds.

What is the role of inconsistency within this view? Inconsistencies are presumably unavoidable. Perhaps it is even natural that our knowledge acquisition leads to inconsistencies. Some cannot be eliminated for the time being, others cannot be overcome except at too high a cost. All this is no reason to give up the search for consistency, or to refrain from believing that some domain is consistent in case we have good reasons to do so. Only if inconsistencies are seen as problematic are we able to properly account for their role as a motor in knowledge acquisition.

NOTES

[1] For an older survey of the domain, see Ayda Arruda "A survey of paraconsistent logic," in A. I. Arruda and R. Chuaqui (eds.), *Non classical Logics, Model Theory and Computability*, 1980, Amsterdam, North-Holland, pp. 3—24; more up to date information will be provided in G. Priest, R. Routley & J. Norman (eds.), *Paraconsistent Logics*, Philosophia Verlag, a book which I hope to be (finally) out by the time this paper appears.

[2] A logic is *strictly* paraconsistent if it does not validate any form of the ex falso quodlibet, e.g., $\sim(A \& B), (A \& B) \vdash C$. See my "Paraconsistent extensional propositional logics," *Logique et Analyse*, 90—91, 1980, pp. 195—234.

[3] Alan Ross Anderson & Nuel D. Belnap, Jr., *Entailment. The Logic of Relevance and Necessity*, vol. 1, Princeton, 1975. Nowhere do they even suggest that true theories might be inconsistent. But they actually subscribe to global paraconsistency, for their systems happen to be paraconsistent, and they hold that one of them, viz. E, is (at least very close to) the correct logic and should be used as the underlying logic of the language in which the systems are described.

[4] G. Priest, *In Contradiction. A Study of the Transconsistent*, Dordrecht, Nijhoff, 1987. All occasional references to pages are to this book.

[5] For a discussion with respect to relevance see my "Two semantically motivated enrichments of relevant logics," in Jerzy Perzanowski, *Essays on Philosophical Logic*, Cracow, Jagiellonian University Press, 1987.

[6] Roman Suszko, "The Fregean axiom and Polish mathematical logic in the 1920s," *Studia Logica*, 36, 1976, pp. 377—380.

[7] J. Kotas & Newton C. A. da Costa, "Some problems on logical matrices and valorizations," in A. I. Arruda, N. C. A. da Costa & A. M. Sette (eds.), *Proceedings of the Third Brazilian Conference on Mathematical Logic*, São Paulo, 1980, pp. 131—146.

[8] Needless to say, there is no specific relation to the definition of truth as correspondence to reality.

[9] Some 'inductive logics' were based on derivability to a certain degree. Actually, they led to lots of *uninteresting* paradoxes, and the whole approach was rejected by Rudolf

Carnap, who devised the most successful 'inductive logic' of our century. See, e.g., his "On rules of acceptance," in I. Lakatos (ed.), *The Problem of Inductive Logic*, Amsterdam, 1968, North-Holland, pp. 146—150.

[10] D. Batens, "A bridge between two-valued and many-valued semantic systems: n-tuple semantics," *Proceedings of the 12th International Symposium on Multiple-Valued Logic*, IEEE, Los Angeles, 1982, pp. 318—322.

[11] If you have doubts about this, remember that the classical metalanguage may contain a paraconsistent negation. In other words, we may retain Priest's paraconsistent negation "$\sim$" and introduce a classical negation "$\neg$" that fulfills both (i) $A, \neg A \vdash B$, and (ii) $\neg A \vdash \sim A$.

[12] The underlying mathematics is not worked out in *In Contradiction*, but if it would be, the system would be able to contain its own semantics. So, if the distinction is expressed by the semantics, it would in principle be expressed by the system itself.

[13] Actually I disagree with this claim. I consider it much more certain that Priest's logic may be consistently described than that rational acceptance and rejection would exclude each other. But then the whole matter is not very central to Priest's book.

[14] I consider such cases in the paper referred to in note 2.

[15] For a nice example, see Larry Laudan, *Progress and its Problems*, Berkeley, University of California Press, 1977.

[16] Any sensible analysis of information should give us $I(A, \Lambda) = 0$ if "Λ" stands for the conjunction of all sentences or something equivalent. There is one caveat: in section 10 it will turn out that "triviality" is not completely unambiguous; to be precise I should say "classical triviality."

[17] No one living after the days of Frege and Hilbert should doubt this.

[18] The proof of this statement occurs in *Entailment*, the book referred to in note 3.

[19] For an example, see pp. 17—18 of R. Routley & R. K. Meyer, "Dialectical logic, classical logic and the consistency of the world," *Studies in Soviet Thought*, 16, 1976, pp. 1—25.

[20] See my "Meaning, acceptance and dialectics," in *Change and Progress in Modern Science*, J. C. Pitt (ed.), Dordrecht, Reidel, 1985, pp. 333—360, and my "Do we need a hierarchical model of science," to appear.

[21] I cannot discuss this here, but I advise the reader to read Katalin Havas's "Dialectic and inconsistency in knowledge acquisition" (this volume) with the discussion about global paraconsistency in mind.

Rijksuniversiteit Gent,
Seminarie voor Logica en,
Wijsbegeerte van de Wetenschappen,
Rozier 44,
B-9000 Gent (Belgium)

DIDERIK BATENS

Addenda to "Against global paraconsistency".

In view of the comments Graham Priest gave me in December 1990, I want to add the following notes to the paper. *Please do not read or copy the paper without the present addition.*

p. 211, 4th paragraph. The two first sentences should be taken literally. I did *not* mean too say that the definition of 'logic' can only be understood from a classical point of view.

p. 212, first paragraph of section 3. Priest does not claim that $0 \in v(A)$ iff $1 \notin v(A)$. According to his terminology, "false" and "not true" do not have the same meaning, and that A is both true and false is not an inconsistency. However, by 'diagonalization' and paradoxes such as the Liar, inconsistencies *do* arise in the semantic metalanguage.

p. 216, after list (2). TA and $\sim FA$ are *not* equivalent. Fortunately, this and the previous mistake do not undermine the arguments made in these sections.

p. 220, statement preceding (1). More accurately: some people believe Disjunctive Syllogism to be correct because it seems to apply in common contexts; there are reasons for this, one of them being the correctness of (1).

last paragraph of p. 220, continuing on p. 221. The important connection stated here is fully independent of the inaccurate wording of the statement preceding (1). Actually, *In Contradiction* contains the possibility to define a strong negation (though not the boolean negation of classical logic), viz. as $\neg A =_{df} A \rightarrow \mathsf{F}$ (where F is the "propositional constant, thought of informally as the conjunction of all formulas", introduced on p. 146).

Diderik Batens