XXI
Reductio ad Absurdum et Modus Tollendo Ponens

G. Priest

1. Introduction*

The existence of true contradictions should be taken seriously. For many reasons we have good grounds for supposing that there are statements of the form $A \wedge \neg A$ which are true. This has been argued elsewhere (see, for example, Priest, 1979) and I do not wish to discuss it now. Rather I wish to discuss the consequences of this for some well-known forms of argument. In particular, if there are true contradictions, both the rules *reductio ad absurdum* and disjunctive syllogism or, to give it its medieval name, *modus tollendo ponens* are problematic. Yet these seem to be an integral part of common reasoning. Should we give them up as invalid? Can we give them up without crippling reason? What exactly is their role in reason? This is the problem I wish to tackle.

2. Logic as an organon of criticism

Logic is conceived very narrowly nowadays as the study of the forms of valid (usually deductive) reasoning. However, traditionally, logic was conceived as a much broader subject encompassing both logic (as now conceived) and rhetoric. Rhetoric is the study of the use of arguments to legitimately convince or persuade (which should not be confused with sophistry, which is the study of illegitimate methods). Thus discussions of begging the question, *ad hominem* arguments, etc. fall in the domain of rhetoric, not logic as modernly conceived. It is a pity that the study of rhetoric has lapsed in modern logic. This is for many reasons, the important one at the moment being that the key to our problem lies there.

Let us start by facing the following objection to paraconsistency. If paraconsistency is correct, then logic is ruined as an organon of criticism. For whenever we attack someone's views by producing an argument against one of them, there is nothing to stop him accepting both his original position *and* our argument, the conclusion of which negates it!

Note first that this objection falls squarely in the domain of rhetoric, for it concerns the use of logic to change the views of another. Secondly, note that what is at issue here is precisely the critical use of *reductio ad absurdum*. When one produces an argument against an opponent, it is useless to start from premisses which he does not accept. This will not convince. One has to start from premisses he already accepts. What the argument then does is show that these premisses, which he believes, entail the negation of something which the person also believes. In other words, one shows that his belief set is inconsistent, and this is supposed to make him change his views. This of course is exactly *reductio ad absurdum*. Finally note that the thrust of the argument against paraconsistency is *not* that if we allow someone to hold on to a contradiction he would be believing the False. (For the paraconsistentist holds that a contradiction may not be the False, and so this objection would beg the question.) Rather it is that if paraconsistency is correct, *reductio ad absurdum* ceases to be an effective organon of criticism—that is, a way of criticizing and thereby changing the beliefs of others.

However, the objection is incorrect and springs from a misconception about how a *reductio* argument works in this context. The assumption it makes is that pressure is put on the person to revise his views when some of them are shown to be false (a contradiction being a manifest falsity). If this were correct then the possibility that the contradiction is true would indeed take pressure off him. However this is not how a *reductio* works. The point of this kind of argument is to get people to change their beliefs, and showing that their beliefs entail a falsity is neither sufficient nor necessary for this. It is not sufficient, since if the person does not *accept* the consequence of his view as false, he will feel under no pressure to change his views. (It is useless showing that someone's views entail the flatness of the earth if the person in question turns out to be a flat-earther!) Moreover it is not necessary, since even if the consequences are indeed true, provided the person disbelieves them he will be under pressure to change his views. Thus we see that the effectiveness of a *reductio* argument depends essentially on showing that a person's beliefs entail things he is wont to reject.

Having got this straight we can now see that a *reductio* argument can be just as effective if paraconsistency is correct. For paraconsistentists are not committed to accepting *all* contradictions. In fact most of us are wont to reject most of them. So in most cases of a *reductio* argument the pressure is still on to change one's belief. Where the paraconsistentist may differ from the next man is that he *may* seriously consider accepting the contradiction and he *may* in the end decide to accept it. (Though in an arbitrary case he is much more likely to reject it.) However this is not a sign of stupidity: it is a sign that he is less narrow-minded and dogmatic than the classicist who rejects the contradiction thoughtlessly and out of hand. In any case the main point has been made. The claim was that *reductio* fails to be an effective form of criticism if paraconsistency is accepted. This has been refuted. An argument by *reductio* may well be effective; it is just not guaranteed to work *every* time.

3. Plato's Euthyphro: an illustration

To illustrate this point it is worth considering briefly Plato's dialogue *Euthyphro*. The *Euthyphro* is a dialogue between Socrates and Euthyphro, who is prosecuting his father for manslaughter. Despite Socrates' doubts, Euthyphro is quite sure that this is the pious course of action. When Socrates asks him to define 'piety' the first definition he produces is that piety is that which is loved by the gods (6e–7a). With characteristic flair Socrates shows that it follows that something may be both pious and impious (8a). Euthyphro later recants. It is easy to read this passage over-simply and think that Euthyphro rejects his definition simply because it implies a contradiction. However, this would be incorrect (see Candlish, 1983). Euthyphro does not balk at the idea of something being both pious and impious (8a). (Neither, for that matter does Socrates who later accepts that something may be both pious and impious (8d)). When, in fact, Euthyphro defends himself against Socrates' attack (8b), he defends himself against the possibility that what he is doing (prosecuting his father) is both pious and impious. Thus he interprets Socrates' argument as showing that he may be doing something impious (even though pious), which he cannot accept. Thus, Socrates' argument is effective without the general assumption that contradictions are unacceptable. It is effective because this particular contradiction is unacceptable.

This illustrates exactly how paraconsistency does not show *reductio ad absurdum* to be a useless form of argument in a rhetorical context. *Reductio* is just a special kind of *ad hominem* argument: an argument which shows certain beliefs to have unpalatable consequences.

4. The general unacceptability of contradictions and the particular acceptability of some

But let us probe deeper. Why, after all, should a paraconsistentist be wont to reject a contradiction? After all, the classical *rationale* for rejecting a contradiction is that it is just plain untrue, and therefore not to be believed. Once we admit that contradictions may be true, why are we wont to reject them?

The answer to this is in part, I think, that once we get used to the idea that some contradictions are true, we may not be quite so wont to reject them. I for one have become more cautious since becoming a paraconsistentist. But this does not get to the root of the problem. For although some contradictions are true, there is still the presumption, and a reasonable one, that an arbitrary contradiction is much more likely not to be true than to be true. For the number of true contradictions is relatively small compared with that of the untrue ones. Of course, "relatively small" is a gloss here. Given any language in which there are sentences that are paradoxical, there will be as many paradoxical sentences as sentences. (If A is paradoxical, so are $A \vee A$, $A \wedge A$, $\neg A$ etc.) However, the frequency with which we meet paradoxical sentences in *ordinary reasoning* is very low. At best (at worst?) paradoxical sentences occur in only certain logico-mathematical, legal and dialectical situations (and maybe a few others). But these hardly account for the subject matter of most of our day-to-day reasoning, and even where they are the subject matter, the frequency of the occurrence of paradoxical sentences is still low. It is precisely the low frequency of paradoxical sentences in our discourse which justifies the general presupposition that a contradiction we meet is not paradoxical. In particular cases we may have to scrutinize this presumption and possibly reject it, but objectively it provides the *rationale* for the effectiveness of *reductio* arguments.

All this may be very well, but it raises the obvious question of how we tell whether a particular contradiction $A \wedge \neg A$ *is* true. The question is well worth asking, but the answer is simple and, almost certainly, disappointing. We find it to be true by finding that A is true and finding that $\neg A$ is true. But how do we find that A is true? Unfortunately there is no universal answer to this. (If there were, life would be much easier!) Every domain of inquiry has its own (fallible) tests for truth. That is all one can usefully say in general. However, it may be worth examining one particular example, the set-theoretic antinomies.

Suppose we start off (as historically we did) with a belief in the axioms of naive set theory. These lead to a contradiction. Should we continue to accept the axioms *and* the contradiction, or reject the contradiction and try to reformulate the axioms? Clearly the thing to do is to investigate both possibilities. Only one has been pursued with much vigour till now, the latter. The results have not been encouraging. The reformulations are always *ad hoc* (with little rationale—at least little that will stand up to gentle pressure), arbitrary (witness how many ways in which it can be done) and have as consequences the rejection of a substantial part of perfectly good, naive reasoning. (Look at the confusions in the foundations of category theory.) Compared with this the other possibility has little that could go wrong. Its only problem is that too much might turn out to be paradoxical. It is all right for Russell's set to be both a member and not a member of itself, but it would be rather too much if we could prove that every ordinal was identical with every other ordinal (as well as distinct from it), or even that $0 \neq 0$. These results seem unlikely. The results of Brady, 1989, are a very important step in settling the issue, but clearly more research needs to be done.

Anyway, once these two possibilities have been investigated we must then choose between the two. The choice is pretty obvious, but we can be a bit more formal about it. As the philosophy of science has shown, it seems virtually impossible to find a unique criterion for theory choice. Rather there is a set of criteria including degrees and amount of *ad hocness*, fruitfulness, problem-solving ability, etc. It is not difficult to see that when the comparison is made, naive, inconsistent, set theory will come way ahead of its competitors. I will leave the detailed case to be spelt out elsewhere (though some of the arguments are given in Priest, 1989, 1983b). Suffice it for the present that I have shown what sort of considerations could lead to the rational acceptance of a contradiction. They turn out to be quite familiar ones.

5. An important point

In the next section I want to turn to the consideration of *reductio ad absurdum* in the context of proof rather than criticism. However, before I do this, there is an important point to be noted. Since we will have two or three occasions to refer to it, it is worth giving it separate consideration.

Suppose someone accepts a disjunction, $A \vee B$. Nothing, as yet, forces him to accept one or other of the disjuncts, A or B. However, let us suppose that he comes to reject one of the disjuncts, say A, but continues to accept $A \vee B$. He then is committed to accepting the other, B. The *rationale* for this is precisely the truth condition for a disjunction: $A \vee B$ is true iff A is true or B is true (or, in another jargon, the primeness of the True). He who commits himself to $A \vee B$ commits himself thereby neither to B nor to A. However he is rationally committed to the acceptance of one of the disjuncts if he rejects the other. Of course the person in question may refuse to be rational and do the right thing; but this is neither here nor there; what he ought, rationally, to do is quite clear. (The point is made forcibly—several times—by Meyer, 1978. See e.g. pp. 87, 90f.)

It may appear to the casual observer that the rationale for this behaviour rests on an application of the disjunctive syllogism: $A \vee B$ is true. Hence A is true or B is true. But A is not true. Hence B is true. However, this is incorrect. The justification for this procedure is a rhetorical one, not a formal one. If $A \vee B$ is accepted, and A is rejected, B ought to be accepted.

If this distinction is not clear, compare it with the following: If someone accepts $A \rightarrow B$ as true and accepts that $\neg B$ is true, then he ought to accept that $\neg A$ is true. The justification for this is the formal validity of *modus tollendo tollens* ($A \rightarrow B$, $\neg B / \neg A$). However if he accepts $A \rightarrow B$ and rejects

B, then clearly he had better reject A. However this is not *modus tollendo tollens.* The crucial point here is that accepting ¬A to be true is different from rejecting A. One can reject A whilst failing to accept ¬A. (If, for example, one holds, perhaps on intuitionistic grounds, that both A and ¬A may fail.) Conversely one can accept ¬A whilst failing to reject A. (If, for example, one countenances the possibility that A is paradoxical.) The disjunctive syllogism argument depends on the acceptance of ¬A, the process we are concerned with hinges on the rejection of A.

To summarize: if someone accepts A ∨ B and rejects A, he ought rationally to accept B. Let us call this the acceptance principle for disjunction (APD). The APD can be symbolized simply. Let us use $\vdash_x$ to mean 'x accepts that' and $\dashv_x$ to mean 'x rejects that' (the notation is Richard Routley's.) Then assuming that x is a rational agent, the APD is simply

$$(\vdash_x A \vee B) \wedge (\dashv_x A) \rightarrow \vdash_x B \qquad \text{(APD)}$$

An obvious question at this point is as follows. We are taking paraconsistency seriously. The APD seems very reasonable to the classical mind, but what if we allow that someone may both accept and reject a claim? Does this under-cut the APD? If x accepts A ∨ B and rejects A, is he still committed to accepting B, even if he accepts A?

There are two important points here. The first is that the paraconsistentist is by no means committed to the view that all contradictions (or pairs of contraries) are realizable. In particular, the pair $\dashv_x A$ and $\vdash_x A$ would not seem to be so. Someone who rejects A cannot simultaneously accept it, any more that a person can simultaneously catch a bus and miss it, or win a game of chess and lose it. If a person is asked whether or not A, he can of course say 'Yes and no'. However this does not show that he both accepts and rejects A. It means that he accepts both A and its negation. Moreover a person can alternate between accepting and rejecting a claim. He can also be undecided as to which to do. But do both he can not. The second point is that even if someone could, *per impossible* both accept and reject A, the APD would not fail. For the worry about the APD hinges on the observation that the fact that someone accepts a disjunction *and* one of the disjuncts in no way entails that they ought rationally to accept the other i.e.

$$\neg((\dashv_x A \vee B) \wedge (\vdash_x A) \rightarrow \vdash_x B)$$

However this does not refute the APD; for both are true. Thus, if the agent rejects A he *would* still be rationally committed to accepting B, even if he could accept A. For two reasons then, the APD is not undercut by paraconsistency.

6. *Reductio ad absurdum* as proof

Let us now move on, to the use of *reductio ad absurdum* in the context of proof. This is a domain with which logicians are much more familiar (though only the formal, not the rhetorical aspects). No longer do we have two people, one of whom is trying to change the views of the other. Instead we have one person, certain of whose views are fixed and not up for review, who then proves others by deducing them from the fixed stock. Traditionally one of the most important such methods of proof has been *reductio ad absurdum*: assume A; deduce from A, and some of the accepted truths, a contradiction; infer ¬A and discharge the assumption A. Perhaps some of the most famous mathematical results have proofs which essentially employ *reductio*: the irrationality of $\sqrt{2}$, the uncountability of the reals, the fact that the power set of x is bigger than x, etc. Yet clearly paraconsistency casts some doubt on the use of *reductio* here. Should we conclude that $\sqrt{2}$ may be rational after all? Let us investigate more closely.

The general form of a proof by *reductio* is as follows:

$$\frac{\Sigma\vdash C \quad \Pi\vdash\neg C}{\Sigma\cup\Pi-\{A\}\vdash\neg A} \qquad (\rho)$$

where '$\Delta\vdash B$' means that there is a derivation of B from all Δ, in which every member of Δ is relevantly used, and $A\in\Sigma\cup\Pi$.

All the other forms of *reductio* are reducible to this, under reasonable assumptions. For example, if $\Sigma\vdash C\wedge\neg C$, then $\Sigma\vdash C$ and $\Sigma\vdash\neg C$. Hence $\Sigma-\{A\}\vdash\neg A$. If $A\in\Sigma$ and $\Sigma\vdash\neg A$ then since $\{A\}\vdash A$, $\Sigma-\{A\}\vdash\neg A$, etc.

Now the existence of true contradictions does not show scheme ρ to be invalid, i.e. we cannot use this fact to construct a counterexample to ρ, and there is no *a priori* reason, therefore, why ρ should not be acceptable in a paraconsistent logic. (It certainly does not lead to the abhorrent *ex falso quodlibet*, provided we take '$\vdash$' seriously. If we were to allow that $\{A, B\}\vdash A$, then we would obtain $\{A, \neg A\}\vdash\neg B$ thus:

$$\frac{\{A\}\vdash A \quad \{\neg A, B\}\vdash\neg A}{\{A, \neg A\}\vdash\neg B}.$$

But there is, in general, no way of bringing it about that $\{A, B\}\vdash A$, as we now know.) Indeed the rule ρ holds in some of the very strong paraconsistent logics such as R.

However ρ certainly fails in some of the weaker ones, such as the logic of my 1980. (Though again certain special cases may hold.) This is not surprising. For although we cannot use the existence of true contradictions to provide a counterexample to ρ, their existence does somewhat undercut the *rationale* for ρ. We can indeed take the main part of a *reductio* proof, the deduction of a contradiction from A and certain axioms, as a criterion of the truth of $\neg A$. But why should we if the contradiction may be true?

Let us suppose then, (as seems likely) that ρ is not generally acceptable paraconsistently. Should we stop using it in proofs? The answer is 'no', and follows simply from our discussion of *reductio* in the context of criticism. For the context of proof is easily assimilated to the context of criticism. X has accepted that all members of $\Sigma \cup \Pi$, are true. Y forces him to admit that $\Sigma \vdash C$ and $\Pi \vdash \neg C$. It is just that X and Y are the same person. Now in this situation X has a choice. All the members of $\Sigma \cup \Pi$, with the exception of A, are not to be touched since they are axioms. X may continue with his acceptance of A, and accept the contradiction $C \wedge \neg C$. However, he may, and with excellent reason (as we have discussed) reject both the contradiction and A. Provided that it is reasonable to suppose that the domain in question is not a paradoxical one (or at least, if it is, that C is not one of the paradoxes), then it is reasonable to reject A.

However, this does not, as yet, give us the formal *reductio* conclusion. X may have rejected A, but, as we have seen, this is not the same thing as accepting $\neg A$. This is where the APD comes in. For if X accepts $A \vee \neg A$ to be true, then when he rejects A he must, by the APD, accept $\neg A$. Now instances of the law of excluded middle may be suspect on some occasions but, by and large, it is reasonable to accept most of its instances. Hence in most cases, the move will be a very reasonable one. Thus we have seen that provided we accept $A \vee \neg A$ and reject the paradoxicality of C, both of which are likely to be highly reasonable, we may accept the conclusion of a formal proof using *reductio ad absurdum.* Crucially the acceptability of the proof rests on the rejection of the possibility that C is paradoxical. This can be put slightly more generally using the notions of local consistency/inconsistency. These are simple but important paraconsistent notions. If Σ is any deductively closed set of sentences, Σ is *locally consistent* at B iff either $B \notin \Sigma$ or $\neg B \notin \Sigma$. Σ is locally inconsistent at B iff it is not locally consistent at B. (The definition can be extended to sets which are not deductively closed via their deductive closures.) A special case of local inconsistency is when Σ is Tr_L, the set of truths of some language L. In many cases (e.g. if L is semantically closed) Tr_L is (globally) inconsistent. However Tr_L may well be locally consistent at some points. Indeed it is obvious that Tr_L is locally consistent at A iff A is not paradoxical. Hence a *reductio* proof, the conclusion of whose subproof is $C \wedge \neg C$, is acceptable provided we may reasonably reject the local inconsistency of Tr_L at C, where L is the language being used.

Let us now return to the question of whether $\sqrt{2}$ is rational, the reals uncountable, etc. The first steps in the analysis of these questions are to a) fix upon the correct system of paraconsistent logic and b) determine whether the specific forms of *reductio* used in these arguments are valid. These are important questions. However, they are too detailed to discuss here. Let us instead assume, at least for the sake of argument, that these proofs do not come out to be formally valid. The question of the acceptability of these results now hinges on the acceptability of a general *reductio* argument, which, as we have seen, hinges crucially on the question of local consistency. No one has ever produced a reason for supposing that the standard second order theory of reals is inconsistent. Hence given any of its members A, it is reasonable to reject the claim that the theory is locally inconsistent at A. Thus it is reasonable to accept that a *reductio* proof works in this context, and therefore that $\sqrt{2}$ is irrational and the reals are uncountable. Turning to the proof that the power set of X is of higher cardinality than X we may well, in virtue of the paradoxes of set theory, have some grounds for suspecting that a paradoxicality occurs in this proof. However, no one has ever made it reasonable to suppose that common or garden sets have inconsistent properties, and so it seems reasonable to accept that the proof shows that for at least most of the sets, X, that mathematicians are interested in, the power set of X is of higher cardinality than X. However, one might have very real doubts about whether the power sets of the Russell set r the set of all ordinals On, and the universal set V, are bigger than the sets themselves.

An objection needs to be rebutted at this point. If we have reason to doubt *reductio* proofs in the locality of r, V and On, are we not endangering the claim that there are true contradictions? For many of the paradoxes are obtained using *reductio* arguments in their vicinity. The answer is 'no'. For a start, Russell's paradox:

$$\forall x(x \in r \leftrightarrow x \notin x)$$
$$r \in r \leftrightarrow r \notin r$$
$$r \in r \wedge r \notin r$$

uses *reductio* only in the very weak form

$$\frac{A \rightarrow \neg A}{\neg A}$$

which is equivalent to the law of excluded middle and holds in most of even the weakest paraconsistent logics. Moreover, the Burali-Forti paradox is proved without using *reductio* at all. For we can give independent arguments for both $On \in On$ and $On \notin On$. If On is the set of (von Neumann) ordinals then On is a well ordered transitive set of ordinals and hence is an ordinal, i.e. $On \in On$. Moreover, since the ordinals are, by definition, well ordered by $\in$, $(\forall \alpha \in On)(\alpha \notin \alpha)$. In particular, therefore, $On \notin On$.

Similarly, consider the paradox of the least indefinable (without parameters) ordinal. Since the collection of ordinals is uncountable, this exists and is indefinable. But we have just defined it. (Note, also that the existence of uncountable ordinals can be proved without a *reductio* proof. Let $\beta = \{\alpha | \alpha$ is a countable ordinal$\}$. β is an ordinal and hence $\beta \notin \beta$. Thus β is not a countable ordinal.) For a further discussion of paradoxes which do not use *reductio* see Priest, 1983a.

Thus the partial rejection of *reductio* proofs does not undercut the case for paraconsistency. However, to summarize this section, we have seen that where local inconsistency is rejected a *reductio* proof is perfectly reasonable. Paraconsistency may well destroy blind faith in *reductio* as a method of proof. But it substitutes instead reasoned belief.

7. *Modus tollendo ponens*

And now let us remain in the context of proof, but turn to a rule of inference which has caused some controversy recently in the circle of relevant logicians. The problem is the acceptability of the inference form *modus tollendo ponens* (MTP), or, to give it a more usual name, disjunctive syllogism. This is the inference form $\{A, \neg A \vee B\} \vdash B$. In 1975, sect. 25, Anderson and Belnap claim that the axiom form of MTP is false, i.e. that $A \wedge (\neg A \vee B) \rightarrow B$ is false. The claim is that this embodies a fallacy of relevance. However, their claim is somewhat less than convincing. For one reason, the antecedent and consequent of this principle *satisfy* Anderson and Belnap's own condition for relevance; that is, they have a variable in common. Admittedly, this was never intended as a sufficient condition, merely a necessary one. However, $A \wedge (\neg A \vee B)$ looks, for all the world, as if it is relevant to B. Moreover, it is very difficult for them to say why we should refrain from using MTP. After all, classically there would appear to be no situation in which the premisses are true and the conclusion false. So why should we not, if we know the premisses to be true, infer the conclusion?

Anderson and Belnap are quite right that MTP is formally invalid (though of course certain of its substitution instances may be valid). However, the explanation of this is provided by paraconsistency. MTP is not a fallacy of relevance: it just plain fails to be truth preserving. If A and $\neg A$ are true, then the premisses are true, whatever B is. Of course, Anderson and Belnap do not believe in true contradictions and so cannot take this line. This makes their rejection of MTP philosophically unstable. It is not therefore surprising to find another relevant logician, Meyer, who also does not take paraconsistency very seriously, arguing that MTP is a perfectly acceptable form of inference (at least in "sane" situations). (See his 1978.) However, in a sense, Meyer is right: MTP *is* acceptable in consistent situations. I have claimed this myself (in my 1979, sect. IV.1). However, the claim should be considered carefully since it is much more subtle than it appears.

It might be thought that provided we add as an extra premiss that $A \wedge \neg A$ is not true, the disjunctive syllogism is valid. Thus let T be the truth connective. Then although we cannot argue from $A \wedge (\neg A \vee B)$ to B, we can argue validly from $\neg T(A \wedge \neg A)$ and $A \wedge (\neg A \vee B)$ to B, i.e. $\{A \wedge (\neg A \vee B), \neg T(A \wedge \neg A)\} \vdash B$. However, this is incorrect. The invalidity of MTP is shown by taking seriously the possibility of true contradictions. Moreover, if the situation is inconsistent, adding an extra premiss to the effect that it is not, will not change the situation. This is noted by Belnap and Dunn (in 1983). In particular, the truth of $\neg T(A \wedge \neg A)$ does not rule out the truth of both A and $\neg A$! If A and $\neg A$ are true, so is $A \wedge \neg A$ and so is $T(A \wedge \neg A)$. But this may be just another true contradiction, and if it is, the antecedents of the inference may be true whilst the conclusion is arbitrary. (Notice that if we accept the T-scheme and contraposition then $\neg T(A \wedge \neg A)$ is *always* true! For $T(A \wedge \neg A) \leftrightarrow A \wedge \neg A$. Hence $\neg(A \wedge \neg A) \leftrightarrow \neg T(A \wedge \neg A)$. But the lefthand side of this is a tautology, classical *and* paraconsistent. See my 1979, III.8.)

Could it be that there is some other condition on A that could be added to the disjunctive syllogism to make it enthymematically valid? The obvious thought is that some rendering of 'A is not paradoxical' ought to do the trick. However this will not do. The trouble here is that there is no *a priori* guarantee that 'A is not paradoxical' is not itself paradoxical. Thus 'A is not paradoxical $\wedge A \wedge \neg A$' may itself be true. And if it is, all of the premisses of the enthymematic disjunctive syllogism are true whilst the conclusion is arbitrary. Nor is it difficult to show that attributions of paradoxicality may themselves be paradoxical. Just consider

(1) This sentence is false and not paradoxical.

If (1) is true it is false and not paradoxical. Hence it is both paradoxical and not paradoxical. If (1) is false then it is either true or paradoxical. If it is true it is both paradoxical and not paradoxical as before. If it is paradoxical, it is certainly true and hence both paradoxical and not paradoxical. Hence (1) is both paradoxical and not paradoxical.

In fact, we can prove that there is no condition on A, C(A), which added to the premiss of a disjunctive syllogism argument, allows us to infer validly the conclusion. (This observation is due to Errol Martin.)
Suppose that for some C(A),

$\{C(A), A \wedge (\neg A \vee B)\} \vdash B$.
Then $\{C(A) \wedge (A \wedge (\neg A \vee B))\} \vdash B$.
But $\{C(A) \wedge A \wedge \neg A\} \vdash C(A) \wedge A \wedge (\neg A \vee B)$
Whence $\{C(A) \wedge A \wedge \neg A\} \vdash B$.

But this *is* a fallacy of relevance. Thus disjunctive syllogism is not to be salvaged by this approach.

To summarize so far: there is no statement that can be made that forces a formula to behave consistently. We can say 'A behaves consistently' but because our "metatheory" is liable to be inconsistent, this cannot *force* consistent behaviour. This is just one of the hard facts of paraconsistent life.

How then is the claim that MTP is acceptable in consistent situations to be understood? It has been suggested that if we withdraw from the domain of actual reasoning to the metatheory, concerning the situation the reasoner is in, we can obtain a correct understanding. Thus let T be some theory (situation, set of sentences etc.). Then given only that T is consistent (i.e. $A \in T \rightarrow \neg A \notin T$,) and prime (i.e. $A \vee B \in T \rightarrow A \in T$ or $B \in T$), we can infer: from $A \in T$ and $\neg A \vee B \in T$ that $B \in T$.(*) This argument is to be found in effect in Routley and Routley, 1972, p. 329.

However, as usual, retreat into the metatheory does not solve the problem but merely relocates it. For the argument for (*) goes essentially as follows:

(1) Suppose $\neg A \vee B \in T$	
(2) Then $\neg A \in T$ or $B \in T$	since T is prime.
(3) But $A \in T$	
(4) Hence $\neg A \notin T$	since T is consistent.
(5) Thus $B \in T$	by (2) and (4).

Now the obvious problem with this argument is that step (5) is an application of MTP. It might be suggested that this opens the Routleys to an *ad hominem* argument. However this would be a mistake. They are, after all, trying to show that *some* applications of MTP are acceptable. However it is clear that the MTP can not be used to justify using the MTP without, in some sense, begging the question. (The situation is exactly analogous to the inductive justification of induction.) Moreover, and the main point in this context, if we are still trying to ascertain under what conditions MTP can legitimately be used, and why, this argument, because of its use of MTP, is of little help to us. The Routleys are indeed aware of the problem here and they do make a swift attempt to justify its application. For they say that ¬A's "non-membership if T can be used to eliminate ¬A's membership *given only a consistent membership relation*". (*op. cit.*, p. 329, my italics). However this will not help us. For it is exactly the claim that MTP is usable in consistent situations whose precise sense and rationale we are trying to determine. We have therefore gone precisely nowhere. (Indeed, if the appeal here is to global consistency we are much worse off: the membership relation is notoriously inconsistent.)

It could be that there are proofs of (*) which do not use MTP. However I know of no such proof and I do not expect to find one. For the sorts of considerations which produced a counterexample to the enthymematic disjunctive syllogism suggest we may find a counterexample to (*). Essentially, all we need is a consistent theory T such that $\neg A \in T$ and $A \in T$ but such that we cannot be content with $B \in T$ for arbitrary B. I leave the production of such a consistent inconsistent theory as an open problem.

Finally, before we leave this approach to MTP let us note that without MTP we can prove a cousin of (*) viz.

$$(T \text{ is consistent } \wedge T \text{ is prime } \wedge (\neg A \vee B \in T) \wedge (A \in T)) \supset B \in T.$$

The proof I leave as an exercise. (However, those who want to cheat can consult Meyer's proof that γ holds for R in the form of a material implication. See his 1978, pp. 59–65.) But this is not a great deal of help. For even given the antecedent we are not able to infer that $B \in T$ unless we can legitimately use MTP, and the *modus operandi* of MTP is that which we are still trying to ascertain.

So far we have drawn a blank. How are we to understand the claim that MTP is acceptable in consistent situations? In fact the answer is very simple and follows easily from our previous discussion of *reductio ad absurdum.* Let us suppose that we have unconditional proofs of A and of $\neg A \vee B$. Then we have an unconditional proof of $A \wedge (\neg A \vee B)$ and hence $(A \wedge \neg A) \vee B$. But as explained in section 4 it is often reasonable to reject a contradiction and so, by the APD, provided we reject $A \wedge \neg A$ we must accept B. This is the solution to the problem of MTP. It is reasonable to accept the conclusion of an application of MTP provided we may reasonably reject a certain inconsistency. But the rejection of $A \wedge \neg A$ is exactly the rejection of the local inconsistency of the context in question at A. Thus all hinges on the rejection of a local inconsistency.

It is absolutely essential to distinguish here between accepting (local) consistency and rejecting (local) inconsistency. It should be obvious by now that these are quite distinct. Moreover, accepting local consistency is not enough to push through MTP. As we have seen, there is no condition on A, the acceptance of which will make MTP work. It is the rejection of (local) inconsistency which does the job.

8. Conclusion: Quasi-valid inferences

My discussion of MTP can be applied to a general class of inferences of which MTP is but one member. These are quasi-valid inferences. An inference is *quasi-valid* if it is classically valid but paraconsistently invalid. (See my 1979, IV.1). Now we can see that the conclusion of a quasi-valid inference is perfectly acceptable provided we may reasonably reject a certain local inconsistency. Consider a quasi-valid inference $A_1 \ldots A_n/B$. Since this is classically valid $\neg(A_1 \wedge \ldots \wedge A_n) \vee B$ is a tautology, classical and paraconsistent. (See my 1979, §III.13). Now suppose we have a proof of $A_1 \ldots A_n$.

Then we can quickly derive a proof of $((A_1 \wedge \ldots \wedge A_n) \wedge \neg(A_1 \wedge \ldots \wedge A_n)) \vee B$. And provided we may reasonably reject local inconsistency at $A_1 \wedge \ldots \wedge A_n$ we can reasonably accept B by the APD, which is that which was to be established.

To sum up the major point of the whole paper, *reductio ad absurdum*, *modus tollendo ponens* and all quasi-valid inferences are perfectly acceptable, provided we can reasonably reject local inconsistency. And this, as we have seen, is usually the case.

Note

* This paper was first read at a symposium on paraconsistent and relevant logic at the Australian National University, April 1980.

References

Anderson, A. R. and Belnap, N. D., Jr., 1975, *Entailment*, Princeton University Press.

Belnap, N. D., Jr. and Dunn, J. M., 1983, "Entailment and the Disjunctive Syllogism", in: G. Fløistad (ed) *Contemporary Philosophy*, vol. 1, Nijhoff.

Brady, R., 1989, "The Non-triviality of Dialectical Set Theory", this volume, chapter XVI, pp. 437–471.

Candlish, S., 1983, "*Euthyphro* 6D–9B and its misinterpretations", *Apeiron* XVII, pp. 28–32.

Meyer, R., 1978, "Why I am not a Relevantist". Research Paper No. 1., Logic Group, Canberra: Australian National University.

Priest, G., 1979, "Logic of paradox", *Journal of Philosophical Logic* 8, pp. 219–241.

- 1980, "Sense, entailment and *modus ponens*", *Journal of Philosophical Logic* 9, pp. 415–435.

- 1983a, "The logical paradoxes and the law of excluded middle", *Philosophical Quarterly* 33, pp. 160–165.

- 1983b, "An antirealist account of mathematical truth", *Synthèse* 57, pp. 49–65.

- 1989, "Classical logic *Aufgehoben*", this volume, chapter IV, pp. 131–148.

Routley, R. and Routley, V., 1972, "The semantics of first degree entailment", *Noûs* 6, pp. 335–358.