40

Logicians Setting Together Contradictories: A Perspective on Relevance, Paraconsistency, and Dialetheism

GRAHAM PRIEST

> You shall never be good logician, that would set together two contradictories: for that, the schoolmen say, God cannot do.
>
> (Thomas Cranmer, cited in the entry for "contradictory" in Little et al., 1973)

1 Introduction

There were doubtlessly many notable features of philosophy in the twentieth century. Perhaps we will have to wait for the perspective afforded by the passage of time to see clearly what they all were. But I think it true to say that one very notable feature is already visible. This is the final breaking of the taboo against inconsistency – the "superstitious dread and veneration in face of contradiction" as Wittgenstein put it (1978: 122). In Western philosophy, since Aristotle onwards, as the quotation from Archbishop Cranmer illustrates, inconsistency has been the ultimate 'no-no.' Accounts of truth, validity, rationality, have all taken it for granted. True, a few enterprising philosophical spirits, notably Hegel, have challenged the orthodoxy. But this was secure whilst its heartland in formal logic lay unchallenged. It is precisely this heartland that was challenged in the twentieth century, and which allowed the unthinkable to become thinkable. The challenge was laid down by paraconsistent formal logics. These logics allow for a discriminating handling of inconsistencies, not the crude 'contradictions entail everything' beloved by the latter-day friends of consistency.

This chapter is not about paraconsistent logics as such. There are many places where readers may go to find out technical details of these logics if they are not already familiar with them. The aim here is to provide a perspective on issues in the philosophy of logic that arise in connection with paraconsistency. This terrain is itself large, though. There is no hope of drawing a comprehensive map – even one of small scale. Rather, readers should regard this essay as a geographical sampler which will (with a bit of luck) encourage them to go and explore the terrain for themselves. Suggestions for further reading are scattered through the chapter at appropriate places. (Short accounts of paraconsistent logics can be found in the articles on paraconsistency in Craig (1999) and Zalta (1999). A much more thorough account can be found in Priest

(2000a), which may be consulted for nearly all the formal details alluded to in this chapter. Another good source of papers on paraconsistency in general is Priest et al. (1989).)

The perspective of the terrain that I will offer here turns around the notion of worlds, actual, possible, and impossible. This will put some order into affairs concerning paraconsistency and two closely connected, but distinct, notions: relevance and dialetheism. I cannot claim that this perspective is a neutral one. On the other hand, I hope that it is a bit more engaging than an account of the kind 'x says this, and y says that.'

Before we start with matters of more substance, let me define some of the crucial notions, so that we know what we are taking about. A propositional logic is *relevant* if, whenever $A \rightarrow B$ is a logical truth, A and B share a propositional parameter, where $\rightarrow$ is the conditional operator. A consequence relation, $\vdash$, is *paraconsistent* if the inference $A, \neg A \vdash B$ (for all A and B) fails. *Dialetheias* are truths of the form $A \wedge \neg A$; and *dialetheism* is the view that there are such things. Let us start with the first of these notions, relevance.

2 Relevant Logic

The thought that for a conditional, $A \rightarrow B$ ('if A then B'), to be true there must be some connection between antecedent and consequent is a very natural one. That is, the antecedent must, in some sense, be relevant to the consequent. The condition is not, of course, satisfied by the material conditionals of classical or intuitionist logic; nor does it appear to be satisfied by the strict conditionals of standard modal logics. In particular, let L be any logical falsehood; then the conditional $L \rightarrow B$ is both materially and strictly valid. Yet, for an arbitrary B, there would seem to be no connection between antecedent and consequent.

Providing an adequate analysis of the notion of the connection is another matter. Even for propositional logics, this is not straightforward. One well-known approach insists that for a conditional to be *really* logically valid it must be logically valid in a truth-preservational sense, and must also satisfy some extra condition of relevance. Thus, we might suggest, $A \rightarrow B$ is logically valid iff $A \rightarrow B$ is a classical tautology (that is, in every interpretation in which A is true, so is B) and, further, $R(A, B)$. Here, $R(A, B)$ is some suitable relationship; for example, that A and B share a propositional parameter. (For explorations of this idea, see the essays in Lehrer and Pollock 1979.)

The notion of conditionality that arises from this approach is a very tractable one, but the approach raises an obvious question. If a conditional is truth-preserving, why is it necessary to add some *extra* condition as well? After all, the whole *point* of a conditional is that its truth provides a guarantee that we can proceed from antecedent to consequent at will. What more than truth-preservation do you need?

A very different approach to relevant logic, that normally associated with the world-semantics of standard relevant logics, regards relevance not as something that should be tacked on to truth-preservation, but as something that falls out of a more adequate notion of truth-preservation. What is wrong with the conditional $L \rightarrow B$, for arbitrary B, is precisely that there are situations in which L holds, but where B does not. For example, let L be the claim that the Peano Arithmetic is complete. This is a logical falsehood. Let B be the claim that Gödel proved that Peano Arithmetic is incomplete. Then the conditional $L \rightarrow B$ is false precisely because there are situations in which Peano Arithmetic is complete and (because of this, indeed) Gödel did not prove its incompleteness.

Situations like this are not logically possible situations. They are logically impossible: logic (and arithmetic) must be different at these worlds. The notion of a physically impossible situation will not raise an eyebrow in these enlightened times. We can all imagine situations where things can accelerate through the speed of light; Newton taught us what such situations might be like. But similarly, we can all imagine situations were the laws of logic are different. We all know what a situation would be like where the law of double negation fails; Brouwer taught us what such situations might be like.

In the simplified world-semantics of relevant logics, logically impossible worlds are normally called *non-normal*, or *irregular*. Their salient feature is that at such worlds conditionals have truth conditions different from those that they have at normal worlds. If w is a normal world, $A \rightarrow B$ is true at w if at all worlds where A is true, so is B. The simplest policy at non-normal worlds is to assign $A \rightarrow B$ an arbitrary truth-value. The rationale for this procedure is straightforward. It is precisely conditionals (or at least, conditionals of this kind) that represent laws of logic. Hence, they should behave differently at worlds where logic is different. How differently? There would seem to be no *a priori* bound on what is logically impossible. Hence a conditional might take on any value. Validity is defined in terms of truth-preservation at *normal* worlds. After all, we want to know what follows from what where logic *isn't* different. (For further details, see Priest 2000b: chapter 9.)

The semantical procedure just described gives a relevant logic. The truth conditions of conditionals at normal worlds are given in terms of truth preservation, but logically valid conditionals are relevant: if $A \rightarrow B$ is logically valid, A and B share a propositional parameter. And this arises because we take into our sweep logically impossible worlds.

The logic obtained in the way that I have described is, in fact, weaker than the logics in the standard family of relevant logics. The stronger logics of the standard family are obtained by evaluating conditionals at non-normal worlds slightly differently. Specifically, an interpretation is furnished with a ternary relation, R; and $A \rightarrow B$ is true at w iff for all x and y such that $Rwxy$, if A is true at x, B is true at y. (See Restall 1993 and Priest 2000b: chapter 10.) What the ternary relation *means* and why one might employ it in this way, is another matter, and one which is still philosophically *sub judice*. (See chapter 38, "Relevance Logic," of this volume for discussion, and for further references to relevant logic.)

Of course, the interpretations of a formal semantics are just abstract sets of certain kinds. They are not themselves the situations about which we reason. (Though we certainly can reason about situations concerning sets.) The sets *represent* situations. What, then, ontologically speaking, are the situations that they represent?

This is a thorny issue, but of a very familiar kind. There are many views concerning what possible worlds are. (See, e.g., the essays in the anthology of Loux 1979.) Some people, such as David Lewis, are realists about them: the worlds are exactly like the one in which we live, but with their own space, time, and causation. For others, such as Stalnaker, they are abstract objects of a certain kind, for example sets of propositions.

For yet others, such as Routley or Sylvan, they are nonexistent objects of a certain kind. The question of impossible worlds adds little, I think, to this debate. Whatever one takes possible worlds to be, impossible worlds are exactly the same kind of thing. Even if one is a realist about worlds, there is no reason, as far as I can see, why impossible worlds could not be of the same kind – worlds just like ours, with concrete individuals in a reality structured by its own space, time, causation, and now we add: logic. It is not even difficult to draw a picture of what such worlds may be like, the art of Maurits Esher often depicts situations where the logically impossible happens (such as geometric objects assuming configurations impossible in Euclidean space). (For a discussion of impossible worlds, see the essays in Detlefsen 1997.)

3 Paraconsistent Logic

The notion of negation, $\neg$, is an important one, and features in many important laws of logic. Negation is a contradictory-forming operator. That is, for any A, one of A and $\neg A$ must be true, and they cannot both be: $\Box(A \vee \neg A)$ and $\Box\neg(A \wedge \neg A)$. These are the logical laws of excluded middle and non-contradiction. Given these laws, in every possible world, $A \vee \neg A$ and $\neg(A \wedge \neg A)$ hold. There will be impossible worlds where, for any given A, $A \vee \neg A$ fails, or $A \wedge \neg A$ holds, though. For exactly this reason the conditionals $B \rightarrow (A \vee \neg A)$ and $(A \wedge \neg A) \rightarrow B$ may fail in relevant logics.

But let us look a little more closely at possible worlds. Given that disjunction behaves normally, the fact that $A \vee \neg A$ holds at such a world entails that either A or $\neg A$ holds. It might be thought that the fact that $\neg(A \wedge \neg A)$ holds at a world entails that one or other of A and $\neg A$ fails; but this does not necessarily follow, even given that conjunction behaves normally. Whether it does depends very much on the truth(-at-a-possible-world) conditions of negation. How negation functions is not at all obvious. In the history of philosophy, many such accounts have been given. According to some, contradictions entail nothing; according to others, contradictions entail everything; and according to yet others, contradictions entail some things but not others. Even in the 20th century, many different formal semantics for negation have been offered.

To see how it may be possible to have all of A, $\neg A$ and $\neg(A \wedge \neg A)$ holding at a world, consider the following very simple semantics. At every world, w:

$\neg A$ is true at w iff A is false at w
$\neg A$ is false at w iff A is true at w

Now, suppose that it is possible for A to be both true and false at a world. Then at that world, both A and $\neg A$ are true. Moreover, given the law of excluded middle, one of A and $\neg A$ is true; so one of A and $\neg A$ is false. Given that conjunction behaves normally, it follows that $A \wedge \neg A$ is false; and so $\neg(A \wedge \neg A)$ is true at the world as well.

Formal semantics where A may be both true and false are not difficult to construct. But it is natural to ask whether there really are possible worlds at which something may be both true and false. This is a fair question. I think it is also a fair answer that the best reasons for thinking this to be possible are also reasons for thinking it to be actual. So let us shelve this question for a moment.

If there are possible worlds at which A and $\neg A$ are true, and validity is defined in terms of truth-preservation at all normal worlds, then the inference $A \wedge \neg A \vdash B$ (*Explosion*) will fail. The notion of consequence delivered will therefore be paraconsistent. Relevant logics are not necessarily paraconsistent. For example, Ackermann's original relevant logic Π' was not. But relevant logics in the standard Anderson–Belnap family are. Conversely, many paraconsistent logics are not relevant (and may also employ a quite different treatment of negation); for example, the da Costa logic C_ω and its like are not.

Given that the inference *Explosion* fails in a logic, it follows that there may be inconsistent but non-trivial theories – that is, sets of sentences closed under logical consequence, which contain A and $\neg A$, for some A, but not every B. Such theories may not be candidates for the truth in any serious sense. They may, as it were, be descriptions of worlds that, though they are possible in a *logical* sense, are clearly very far from the actual world. Recall, after all, that even consistent worlds where frogs turn into people, and rich capitalists all give their money to the poor, are logically possible.

For all that, these theories may yet be important and interesting; and this is so for many reasons. For a start, such theories can be *mathematically* interesting. They may have a significant abstract structure which demands mathematical investigation, just as much as consistent ones do. (After all, one does not have to be an intuitionist to find intuitionist structures mathematically interesting.) Thus we have the rapidly developing study of inconsistent mathematical structures, a notable example of which are inconsistent arithmetics. (For an introduction to the whole area of inconsistent mathematics, see Mortensen 1995.)

Inconsistent theories may have physical importance too. An inconsistent theory, if the inconsistencies are quarantined, may yet have accurate empirical consequences in some domain. That is, its predictions in some observable realm may be highly accurate. If one is an instrumentalist, one needs no other justification for using the theory. And even if one is a realist, one may take the theory, though false, to be a significant *approximation* to the truth. This would seem to be how those who worked on early quantum mechanical models of the atom regarded the Bohr theory, for example. The theory was certainly inconsistent, as all agreed; yet its empirical predictions were spectacularly successful.

Finally, inconsistent theories may have practical importance too. This would be the case if our best understanding of how a piece of technology functions were provided by an inconsistent physical or mathematical theory of the kind we have just considered. Perhaps more importantly at the present, in information-processing of a kind that is now essential to everyone's life, there is always the possibility, indeed the high probability, of information that is inaccurate; inaccurate to the point of inconsistency. Where we discover that our information is inaccurate we will, of course, want to correct it. But on many occasions we may not know that it is inaccurate; nor may there even be a practical way of finding out. There is no algorithm, after all, for determining when information expressed in the language of first-order logic is inconsistent. In such circumstances, employing a paraconsistent logic is the only sensible strategy. We do not want our information-processor to tell us that the quickest way from Brisbane to Sydney is via New York, just because it has corrupt information about bus times in Moscow.

Before we leave the issue of paraconsistency as such, let us return to the Bohr Theory of the atom. A major reason why this was never regarded as a serious candidate for the truth was not so much that it was inconsistent as that it refused to allow inferences that were obviously truth-preserving, on pain of empirical inadequacy. In particular, it refused to allow the inference of adjunction: $A, B \vdash A \wedge B$. This was because the theory was chunked in a certain sense. The theoretical postulates were formed into certain groups (not necessarily disjoint). In computing the stationary states of the atom the quantum postulate was employed, but not Maxwell's electrodynamic axioms. In computing the results of transitions between the stationary states, Maxwell's axioms were employed. Within each chunk inference was allowed free reign. There was also a limited amount of information which was allowed to permeate between the chunks; but what one was not allowed to do was to take arbitrary information, A, from one chunk, and add it to another, containing the information B, and so infer the conjunction $A \wedge B$. (See Brown 1993.)

The chunking strategy is one that is employed in certain kinds of paraconsistent logics of the non-adjunctive variety. Specifically, given inconsistent premises including A and $\neg A$, one is not allowed to put these together in the same chunk to infer $A \wedge \neg A$, and so an arbitrary B, classical logic being the logic standardly in force in each chunk.

There are many ways of enforcing the chunking strategy, but the various details need not concern us here. I want merely to note that the strategy has no intrinsic connection with paraconsistency. For a start, there may be reasons for chunking information that have nothing, as such, to do with inconsistency. For example, one might chunk, not because failure to do so would lead to contradiction, but simply because failure to do so would lead to empirical inadequacy: false observational predictions. Or one may want to keep the information obtained from different sources in different chunks, not because the chunks may be mutually inconsistent (though they may be); but because information sources, such as witnesses, are notoriously unreliable. The fact that the same information occurs in different chunks speaks to its reliability, and is therefore itself a significant piece of information.

Moreover, and most importantly, there is no reason why the logic in force in each chunk must be classical logic. It could itself be a paraconsistent logic. For example, suppose that one of the sources of information was dialetheic, endorsing certain contradictions (though not all). In this case, to determine the proper content of that chunk, one would need a paraconsistent logic. Chunking strategies can, in fact, be employed with any kind of logic within the chunks – even with different logics within different chunks.

4 Dialetheism

Let us come back to worlds again. Someone may well hold that there are possible worlds that are inconsistent without holding that the *actual* world is. After all, the actual world is special. Truth at that world coincides with truth *simpliciter*. And truth has special properties all of its own. For example, one might well hold that for any A, $\neg A$ is true iff *A fails to be true*, whilst this is not true of worlds in general. The claim that the actual world is inconsistent, though, is dialetheism. What reasons, then, are there for supposing that some contradictions are true?

There are many such reasons. (A number are discussed in Priest 1987.) Perhaps the best concerns the paradoxes of self-reference. One of the oldest, and most notorious, of these is the liar. This is a sentence, L, of the form $\neg T\langle L\rangle$, where T is the truth predicate, and angle brackets represent some naming device. The T-schema, $T\langle A\rangle \leftrightarrow A$ (for any sentence A), is an intuitively correct principle about truth. Substituting L in this gives $T\langle L\rangle \leftrightarrow \neg T\langle L\rangle$; and contradiction is but a few logical moves away.

The liar paradox and self-referential arguments of its kind, like Russell's paradox, are apparently sound arguments ending in contradiction. Of course, many other paradoxes are this too. But it is a striking fact about the paradoxes of self-reference that, though they have been the centre of so much philosophical attention for over 2000 years (at least the older ones), there is no consensus as to what, if anything, is wrong with them.

There are also reasons for supposing that the failure to solve the paradoxes is not simply a matter of lack of skill on the part of logicians. The paradoxes seem enormously robust. When steps are put forward to solve them, the contradictions concerned just seem to move elsewhere (in the shape of so called 'strengthened paradoxes'). It seems that contradiction is inherent in the various set-ups, and that all we can do is juggle it around. It is like those old-fashioned children's puzzles where one moves around pieces inside a frame, to try to achieve some predetermined pattern. Given a space in the frame, any adjacent piece may be moved into it. In this way, one can fill any given space; but filling it always creates another. There is always a space somewhere.

The appearance of the inevitability of contradictions is, I think, correct. The contradictions involved in the paradoxes of self-reference are, in a sense, inherent in thought. Our conceptual structures give us, at once, mechanisms for totalization and mechanisms that provide the ability to break out of any totality, such as diagonalization. The two mechanisms together produce contradiction. (This theme is explored at length in Priest 1995.) If this is the case, then certain contradictions are not only *actually* true, but, being inherent in thought, are *necessarily* true.

Of the other *prima facie* examples of dialetheias that one might cite, let us look at just one more. Boundaries are very puzzling things. They are almost contradictory objects by definition. For they both separate and join the areas of which they are the boundary. It is not, perhaps, surprising, then, that various kinds of boundaries seem to realize contradictions. Consider, for example, the boundary between the interior of a room (that which is in it) and the exterior (that which is not in it). If something is located on that boundary, is it in the room or not in it? Or suppose that a radioactive atom instantaneously and spontaneously decays. At the instant of decay, is the atom integral or is it not? In both of these cases, and others like them, the law of excluded middle tells us that it is one or the other. Yet the boundary is symmetrically placed with respect to each of its sides; so the only possibility that Reason countenances is a symmetric one. Thus, the object on the boundary of the room is both in it and not in it; and the atom at the point of decay is both integral and non-integral.

We see, then, there are reasons, at least *prima facie* reasons, for supposing that there are dialetheias. What reasons are there for holding such conclusions to be mistaken; that is, for holding that for no A are A and $\neg A$ both true?

The classical defense of this view is to be found in Aristotle's *Metaphysics*, Γ, 4; but this is hardly very successful. The major argument in the chapter is tangled and convoluted. It is not clear *how* it is meant to work, let alone *that* it works. The other arguments are short and similarly unsuccessful. Many of them do not even get to first base, since their conclusion is patently that it is not the case that *every* contradiction is true – or even that it is not possible to believe that every contradiction is true – things which are quite compatible with some contradictions being true. Moreover, I know of no way of reworking any of these arguments which makes them successful. (For a discussion of all of this, see Priest 1998.)

It is a singular fact that no philosopher since Aristotle has attempted a sustained defence of the view. What arguments are there? Here are a couple of notable ones. The first starts from the claim that for any statement to be meaningful, it must exclude something: it must say that we are in *this* situation, rather than *that*. But, the argument continues, the negation of a sentence holds in exactly those situations that the sentence does not hold in. Hence, we cannot have both A and $\neg A$ holding at any situation, and in particular, in the actual situation.

The argument appeals to a contentious theory of negation, one that a paraconsistent logician is likely to dispute. But let us suppose, for the sake of the argument, that the theory can be substantiated. The argument still fails. The claim that a meaningful sentence must exclude something, the other of its major premises, is precisely not available to classical logicians. For according to them, all necessary truths hold at *all* worlds. In particular, given the account of negation, the claim that $\forall A \neg(A \wedge \neg A)$, since it holds in all worlds, is itself meaningless! Ironically, it is the broader, relevant/paraconsistent, perspective that can accommodate the view about meaning in question. For given that there are impossible worlds, all claims, even logical truths, fail at some world.

A second argument appeals to the fact that we never observe contradictory situations: we never see a person both sitting and not sitting; we never see a group of people in which there are both three and not three. (Even if contradictions arise at instantaneous transition states, being instantaneous, these are not observable.) So there is good reason to believe that contradictions are never true. The argument is an inductive one, which might be thought strange, since the conclusion is supposed to be a logical truth; but one can collect *a posteriori* evidence for *a priori* truths: for example, it is *a priori* true that if a is taller than b, and b is taller than c, then a is taller than c; and we can collect evidence for this by going around measuring lots of as, bs, and cs.

The argument is not just an inductive one, though: it is not a very good inductive one. For the crucial question is whether the sample from which we are inducing is, in fact, a typical one; and the observable realm is not very typical in many ways. This is one of the lessons of modern science. Unobservable realms, particularly the micro-realm, behave in a very strange way, events at one place instantaneously affecting events at others in remote locations. Indeed, it would sometimes (in the well-known two slit experiment) appear to be the case that particles behave in a contradictory fashion, going through two distinct slits simultaneously. The micro-realm is so different from the macro-realm that there is no reason to suppose that what holds of the second will hold of the first. *A fortiori* when we move away from empirical realms altogether, the realm of sets appears to be inconsistent. Why should the way that observable things behave tell us anything about this?

Giving arguments to the effect that A and $\neg A$ are *never* true together is clearly a difficult matter. Some have concluded that it is impossible: this fact is so basic that there is no way that one can argue for it at all – at least, without begging the question. Despite the fact that Aristotle did give arguments, this was, in fact, his view of the matter. Only the 'uneducated' would ask for a proof (*Metaphysics*, $1006^a 5$–7). Whether his own views were consistent on this matter we will leave Aristotle scholars to argue about! The two arguments we have just looked at show at least the *possibility* of mounting sensible arguments for the claim. And though the arguments do not work, they do not fail simply because they beg the question.

5 Boolean Negation

At this point, let us look at a more subtle objection to dialetheism. This starts by conceding that the truth-in-a-world conditions of negation may well be what they are claimed to be by a paraconsistent logician. But, it continues, we can characterize a connective, call it ∇, by *giving* it the classical truth conditions. For every world, w:

BN ∇A is true at w iff A is **not** true at w

(And maybe giving it appropriate falsity conditions too.) I have boldfaced the negation in the conditions so that we can keep track of it in what follows. Since ∇ has the truth conditions of classical negation, it satisfies all the inferential principles we associate with that connective. In this context, ∇ is usually called *Boolean negation*, and contrasted with some relevant/paraconsistent negation (*RP negation*). Whether or not it is $\neg$ or ∇ that expresses vernacular negation is now largely irrelevant. For what the classical logician *wished* to express by negation can be expressed by ∇.

Now it would certainly appear to be the case that we can characterize a connective with the truth conditions BN. The problem is in establishing that this connective really does have all the properties of classical negation. To establish, for example, that $A, \nabla A \vdash B$ we have to reason: (1) for any w, it is **not** the case that both A and ∇A hold at w; hence (2) for any world, w, if A and ∇A hold at w, so does B. But what is this **not**? If it is $\neg$, the last inference is clearly invalid. ($\neg C \vdash C \rightarrow D$ is not valid in any paraconsistent logic.) Suppose, then, that it is ∇. If ∇ satisfies all the properties of classical negation, then (2) is acceptable. But recall that we were precisely in the process of mounting an *argument* that it does have these properties. Such a claim therefore simply begs the question.

It is sometimes suggested that metatheoretic truth-conditions of the kind BN are always given employing classical logic – in which case the inference in question is valid. But metatheory is not necessarily classical. For example, intuitionistic metatheory of intuitionistic logic is well-known. (See, e.g. Dummett 1977: chapter 5.) And why, in the last instance, if you think that one particular logic is correct, should there be any significance to a metatheory for it couched in a different, and incorrect, logic?

For a paraconsistent logician, the connective whose truth conditions are given by BN is a perfectly sensible connective. It just doesn't satisfy the classical advertising hype that goes with it. Could we not, though, simply stipulate that ∇ is a connective whose meaning is determined by the proof-theoretic rules of classical negation? In a gem of an article, Prior (1960) pointed out that one cannot simply lay down a set of rules and expect it to characterize a meaningful connective. Suppose that we try to extend our set of logical operators by adding a new binary connective, * (*tonk*), satisfying the rules $A \vdash A * B$ and $A * B \vdash B$. Then all hell breaks loose: we can infer everything. If * were a meaningful connective, its addition would not interfere with the pre-existing machinery. In particular, then, its addition would not allow us to infer any sentence not containing * that was not inferable before. In technical jargon, the extension by *tonk* would be *conservative*. The fact that the extension is not conservative shows, therefore, that *tonk* is not meaningful.

Now, in a similar way, suppose that ∇ were stipulated to satisfy all the inferential principles of classical logic. Then given machinery which includes the *T*-schema and self-reference, we could construct a sentence, *L*, of the form $\nabla T\langle L\rangle$, and all hell would break loose in the same way: everything could be inferred. Hence, its addition is not conservative; so no meaningful connective can satisfy all the principles of classical negation. (It does not follow that there are not operators that behave as classical negation does in limited contexts. In situations that are consistent $\neg$ behaves in exactly that way.)

Of course, the question of conservative extension is relative to what one is extending. In the argument concerning ∇, one is extending machinery that is broader than propositional or first-order logic. But to restrict one's logical machinery to just this is somewhat arbitrary. The truth predicate, governed by the *T*-schema, would seem to be just as much a logical constant as the identity predicate, governed by its usual axioms.

It may be something of a shock that Boolean negation is meaningless. But what is, and what is not, a meaningful specification is not a matter of self-evidence. Such questions are highly theory-laden. And a dialetheist about the paradoxes of self-reference lines up with an intuitionist on this front. For the intuitionist, too, Boolean negation is meaningless, though for quite different reasons. (For an intuitionist, it must be possible, in principle, to recognize the truth of any sentence. Sentences starting with a Boolean negation do not have this property.)

There is an illuminating argument to the effect that Boolean negation is indeed meaningful, which goes as follows. (A version of this can be found in Batens 1990.) It must be possible to deny something, that is, to indicate that one does not accept it. Even dialetheists, after all, need to show that they don't accept that $1 = 0$. Now, if $\neg A$ is compatible with *A*, then asserting $\neg A$ cannot constitute a denial. To deny *A* one must assert something that is incompatible with it; so Boolean negation must make sense. We need to assert something with this force in denying.

Now, denial is a certain kind of illocutory act, an act with a certain linguistic force. It conveys the information that the utterer does not accept the thing denied. Other kinds of linguistic force include: asserting, questioning, commanding. Since Frege, it has been common to hold that denying is not an act *sui generis*. To deny *A* is simply to assert its negation. But this cannot be right. For example, we all, from time to time, discover that our views are, unwittingly, inconsistent. A series of questions prompts us to assert both *A* and $\neg A$ for some *A*. Is the second assertion a denial of *A*? Not at all; it is conveying the information that one accepts that $\neg A$, not that one does not accept *A*. One does this as well.

Denial, then is a linguistic act *sui generis*. Moreover, from the fact that one can deny A, it does not follow that there is some operator on content, ∇, such that to deny A is to assert ∇A, any more than from the fact that one can command that A it follows that there is some operator on content, !, such that to command A is to assert $!A$. Linguistic force is an element of communication *over and above* content. Suppose I utter 'The door is open'; then depending on the context, this could be an assertion, a question, a command. Similarly, if I utter 'It is not the case that A', this could be an assertion of $\neg A$, a denial of A – or even a command, or an act with some other linguistic force. The question is simply one of whether the act is intended to convey the information that the speaker does not accept A, or something else. Denial, then, is a linguistic act, performed by dialetheist and non-dialetheist alike, which in no way presupposes the meaningfulness of Boolean negation. (For further discussion of the material in this section, and negation in general, see Priest 1999.)

6 The Logical Choice

The issues that we have been dealing with concern, either implicitly or explicitly, the question of what the correct logic is. And this raises the question: how do you decide this matter? How, for example, does one determine the correct truth-at-a-world-conditions for negation?

Some have thought that such questions are silly. Logical principles are *a priori* obvious. Those who deny them are uneducated or insane. Such a view could be held, however, only by someone largely ignorant of the history of logic. In the history of logic there are dozens of different accounts of how negation functions, of when a conditional is true, of what inferences are valid – and corresponding disputes. (For a good discussion, see Sylvan 2000.) Moreover, views that have been well-entrenched for centuries have been overturned. For hundreds of years, 'All *A*s are *B*s' was held to entail 'Some *A*'s are *B*'s,' though it is not now. It may well have been the case that some of these principles were thought to be obvious. What was obviously true to one person, may be obviously false to another.

Such questions must, then, be taken seriously. But how do you resolve disputes about the correctness of logical principles themselves? Such disputes are liable to invoke arguments of a form whose very validity is itself disputed.

In disputes that involve high-level and very abstract principles, such as disputes about logic, it is not to be expected that any individual and simple argument, even if its validity is agreed upon by both parties, will be decisive. Arguments of any complexity invoke sundry 'auxiliary assumptions,' which may always be questioned. One is always, therefore, looking at package deals – theoretical complexes that have to be evaluated as a whole. In the case of logic, the package is liable to spread beyond principles simply about validity. There is such an intimate connection between truth and validity, for example, that questions about the nature of truth are likely to be embroiled in the debate as well.

How does one assess such a complex, then? First of all, theories are always proposed to account for some phenomenon, to explain some data; and the first consideration is always how adequate an explanation is provided. In the case of logic, we have intuitions about which inferences are valid and which aren't; which conditionals are true and which aren't; and so on. We must look to see how well the theory accounts for the data. If a theory gives a result that is at variance with them, this is not fatal, but at least we must be able to explain the incongruity. For example, in virtually all relevant and paraconsistent logics, the disjunctive syllogism (A, $\neg A \vee B \vdash B$) is invalid. If we have an intuition that the Syllogism is valid, or at least that it is correct to use it on certain occasions, we must explain why this is so. We may say, as many have said, that the Syllogism is acceptable provided that we are reasoning about a consistent domain – just as an intuitionist may apply the law of excluded middle provided that they are reasoning about finite domains.

Adequacy to the data is not, therefore, likely to be a definitive factor. We have to invoke other criteria. The question of what these criteria are leads to well-known debates in the philosophy of science. Possible candidates include the following: the less a theory invokes *ad hoc* hypotheses the better it is; the more it gives a unified account of its subject matter, the better it is; the more a theory leads to new conceptual developments (fruitfulness), the better it is. There may be many other criteria too. For example, the first two criteria I just mentioned fall under the banner of simplicity; there may be other criteria that fall under this banner too.

Is inconsistency a negative criterion? If the logic of the theoretical complex is explosive, then everything will follow, and this is going to play havoc with the adequacy of the theory to handle the data. So inconsistency is highly relevant. If a paraconsistent logic is used, though, this is not necessarily going to be the case. Is consistency, in this case, a *sui generis* criterion? Is it the case that a theory that is more consistent than another is *ipso facto* a better theory? This is a question that cannot be divorced from the *rationale* for epistemic criteria; and this is a notoriously difficult question. Why, for example should simplicity of any given kind be a positive criterion? If there is some reason for supposing that reality is, quite generally, very consistent – say some sort of transcendental argument – then inconsistency is clearly a negative criterion. If not, then perhaps not.

Let me illustrate some of the preceding points concerning theory-choice. Suppose, for example, that one is comparing classical logic and a paraconsistent logic, as providing accounts of validity for sentences concerning truth-functional operators. As I noted, one cannot simply close ones eyes to other things. The *T*-schema and the inferences that this permits also strike us as valid. If classical logic is correct (and self-reference is legitimate), then this cannot be so: triviality is only a few steps away. Hence, some account of truth must be given which explains away the *T*-schema. If one accepts an appropriate paraconsistent logic, however, one can endorse a natural and simple account of truth: truth just is that notion characterized by the *T*-schema. We must compare, therefore, a package deal concerning (at least) Logic + Truth. Now, most paraconsistent logics are more complex than classical logic – though perhaps not much more so in the simplest cases. But all consistent accounts of truth are enormously more complex than the natural account, involving infinite hierarchies, epicycles to avoid strengthened paradoxes, and so on. What, then, is the simplest overall package? I leave you to judge.

The preceding discussion of theory-choice is, of course, quite general. Though I have couched it in terms of choice of logic, it applies just as much to a choice of any other kind of theory. In particular, it shows how it may be rational to accept an inconsistent theory. (Paraconsistent logic plus the *T*-schema and self-reference is, indeed, inconsistent.) Even if inconsistency is a negative criterion, simplicity and consistency may well pull in opposite directions; a high degree of simplicity may outweigh a low degree of inconsistency. The discussion also shows something else. It is often claimed that if it could be rational to accept a contradiction, a person could never be forced, rationally, to give up any view. For there is nothing to stop the person accepting both their original view and the objection put to it, which is inconsistent with it. It is clear now that this objection fails. It is rational to give up a theory if there is a better one. And even if one can rationally accept an inconsistent theory (or theory plus objection) this may be trumped by a position that is simpler or has greater epistemic virtue of some other kind.

7 Conclusion

God, according to Cranmer in the quote with which we started, cannot set two contradictories together. Cranmer, Archbishop though he was, sold God short (though it was not this for which he was burned at the stake): contradictories can be set together by much lesser creatures. In the last 60 years, logicians have been setting them together in many ways. They may set them together in impossible worlds, to give relevant logics, logics which provide accounts of the conditional which make other accounts look crude and undiscriminating. They may set them together in possible worlds, to provide paraconsistent logics, logics which allow for the sensible handling of inconsistent information and theories. Or if they are daring, they may set contradictories together in the actual world, to allow for things such as a simple and natural theory of truth. These developments in logic, like all interesting new developments, are contentious. And no doubt the issues flagged in this essay will continue to be debated in the foreseeable future. So will many related questions: for the logical views that we have been discussing have implications that spread through metaphysics, epistemology, and many other areas of philosophy. One may presently only speculate as to what lands there are on the far side of the terrain I have been mapping.

References

Batens, D. (1990) Against global paraconsistency. *Studies in Soviet Thought*, 39, 209–29.

Brown, B. (1993) Old quantum theory: a paraconsistent approach. *Proceedings of the Philosophy of Science Association*, 2, 397–441.

Craig, E. (ed.) (1999) *Routledge Encyclopedia of Philosophy*. London: Routledge.

Detlefsen, M. (ed.) (1997) *Notre Dame Journal of Formal Logic*. 38, no. 4 (a special issue on impossible worlds).

Dummett, M. (1977) *Elements of Intuitionism*. Oxford: Oxford University Press.

Lehrer, K. and Pollock, J. (1979) *Philosophical Studies*, 26, no. 2 (a special issue on relatedness logics).

Little, W., et al. (1973) *The Shorter Oxford English Dictionary*, 3rd edn. Oxford: Oxford University Press.

Loux, M. J. (1979) *The Possible and the Actual: Readings in the Metaphysics of Modality*. Ithaca, NY: Cornell University Press.

Mortensen, C. (1995) *Inconsistent Mathematics*. Dordrecht: Kluwer Academic.

Priest, G. (1987) *In Contradiction: A Study of the Transconsistent*. The Hague: Martinus Nijhoff.

Priest, G. (1995) *Beyond the Limits of Thought*. Cambridge: Cambridge University Press.

Priest, G. (1998) To be and not to be – that is the answer. On Aristotle on the law of non-contradiction. *Philosophiegeschichte und Logische Analyse*, 1, 91–130.

Priest, G. (1999) What not? A defence of a dialetheic theory of negation. In D. Gabbay and H. Wansing (eds.), *What is Negation?* (pp. 101–20). Dordrecht: Kluwer Academic.

Priest, G. (2000a) Paraconsistent logic. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, 2nd edn. Dordrecht: Kluwer Academic.

Priest, G. (2000b) *Introduction to Non-classical Logic*. Cambridge: Cambridge University Press.

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

Prior, A. (1960) The runabout inference ticket. *Analysis*, 21, 38–9.

Restall, G. (1993) Simplified semantics for relevant logics (and some of their rivals). *Journal of Philosophical Logic*, 22, 481–511.

Sylvan, R. (2000) A preliminary western history of sociative logics. In D. Hyde and G. Priest (eds.), *Sociative Logics and their Applications* (ch. 5). Aldershot: Ashgate Publishers.

Wittgenstein, L. (1978) *Remarks on the Foundations of Mathematics*, 3rd edn. Oxford: Basil Blackwell.

Zalta, E. (ed.) (1999) Stanford internet encyclopedia of philosophy. http://plato.stanford.edu/contents.html.