Spiking the Field-Artillery

Graham Priest

I submit my cause to the judgment of Rome.
But if you kill me, I shall rise from my tomb
To submit my cause before God's throne.—Thomas Becket.[1]

1 Introduction

Deflationism about truth centres upon the idea that, in some sense, there is no more (or less) to the claim that $\langle A \rangle$ is true than there is to A itself. (Angle brackets indicate some appropriate name-forming device here.) Clearly, it is difficult to hear this view in a way that does not endorse Tarski's T-schema in full generality. But if one does this, then, in the natural course of events, contradiction arises in the form of the Liar and similar paradoxes. Deflationists have often prevaricated over this matter, suggesting the imposition of *ad hoc* restrictions on the T-schema.[2] The unsatisfactoriness of this is clear, however. It is much more natural to accept the full T-schema and the contradictions to which this gives rise, but to use a paraconsistent logic which isolates the paradoxical contradictions.[3]

In 'Is the Liar Both True and False?'[4] Field provides another way in which

[1]T. S. Elliot, *Murder in the Cathedral*, II: 198-200.

[2]See, e.g., Horwich (1990), p. 41.

[3]See Armour-Garb and Beall (2003) and Beall (200a). Not that dialethism about the paradoxes is committed to deflationism about truth: dialetheism goes with any theory of truth. See Priest (2000).

[4]Field (200+). In what follows, page and section references refer to this unless otherwise indicated. All italics in quotations are original.

the T-schema may be accommodated, but this time consistently. He does this by proposing a logic with a novel sort of conditional, to be employed in formulating the T-schema, but without the Law of Excluded Middle (LEM). Field takes it that his construction provides a better solution to paradoxes such as the Liar than any consistent proposal currently on the market. Here I agree. He also argues that it is better than a dialetheic solution. Indeed, he suggests that his construction entirely undercuts dialetheism. Here I disagree. This paper explains why.[5]

2 Preliminary Matters

Field's paper starts with a number of worries about dialetheism—or at least the way that I have formulated it—and though they are not central to the thrust of his criticism, they are not completely unconnected either, so let me start by taking them up.

Let $\bot$ be a logical constant such that, for all A, $\bot \vdash A$.[6] Classically, $\bot$ is equivalent to $B \wedge \neg B$ (for any B). Field thinks that should these two notions come apart (though he does not, himself, think that they do), it is preferable to use the term 'contradiction' for something equivalent to the former; I use it for the latter. As he points out, this is a terminological matter, and nothing hangs on it. For my part, I simply note that history is on my side in this usage. Traditionally, negation has been thought of as a contradictory-forming functor, so that A and $\neg A$ are contradictories, and pairs of the form A and $\neg A$ a contradiction. (The conjunction is not an issue here, as Field points out.) Moreover, authorities from Aristotle to Hegel, *none of whom subscribed to Explosion*, called them this.

Next, a dialetheist holds that the liar sentence, L, is both true and false (i.e., has a true negation): $T\langle L\rangle \wedge T\langle \neg L\rangle$. If one subscribes to the principle:

(*) $T\langle \neg A\rangle \leftrightarrow \neg T\langle A\rangle$

[5]I'm grateful to a number of people for illuminating discussions concerning the matters in this paper; principally to Hartry Field himself, but also to Brad Armour-Garb, Allen Hazen, and Greg Restall.

[6]See Priest (1987), 8.5.

it also follows that L is not true and not false: $\neg T\langle L\rangle$ and $\neg T\langle\neg L\rangle$. Thus we have, $T\langle L\rangle \wedge T\langle\neg L\rangle \wedge \neg T\langle L\rangle \wedge \neg T\langle\neg L\rangle$. It seems 'misleading', says Field, to characterise the status of the Liar simply by the first two conjuncts. Moreover, given that $\neg(A \wedge \neg A)$ as well, it also follows by a bit of juggling that, for any A, $\neg(T\langle A\rangle \wedge T\langle\neg A\rangle)$. This is the negation of dialetheism, the view that some contradictories are both true. If one is a dialetheist, this does not prevent dialetheism being true too; but, thinks Field, it is better to define dialetheism in such a way that the view is not an inherently contradictory one.

Now, as a matter of fact, I don't subscribe to (*);[7] but even if I did, I don't feel the force of the objection that Field feels. It seems to me to be perfectly natural to characterise the Liar as both true and false, since—given (*)—this entail its other properties. And it would be most misleading to characterise it by some other pair of conjuncts with the same property. Thus, for example, given that there are consistent truth-value gap solutions to the Liar, to describe it as neither true nor false would be *most* misleading. As for dialetheism itself being inconsistent, my aim has never been to avoid inconsistency, but to tame it. I don't think that the contradictory nature of dialetheism is worse than any of the other contradictions I subscribe to.[8]

Field then raises a more substantial objection to dialetheism. The point is a familiar one. How can criticism of a view be possible if it is open to a person simply to accept their view together with the content of any objection that is put to it? The answer, however, is simple. Accepting A and $\neg A$ is always a move in *logical* space. This is so even if one accepts classical logic. After all, the person who accepts everything accepts classical logic. This does not mean that it can be done *rationally.* In a nutshell, it is rational to accept a view if it comes out better than any rival on the weighted sum of good-making criteria—such as ontological leanness, simplicity, non-(*ad hoc*ness), maybe even consistency, etc. It is rational to reject it if a rival comes out better.[9] In particular, the view of a person who accepts something or other

[7]See Priest (1987), 4.9.

[8]See Priest (1987), p. 91.

[9]This is discussed in Priest (1987), 7.5, and in greater detail in Priest (2001a).

plus the content of an objection put to it may well fail this test.[10]

3 Revenge is Sweet

With these preliminary points out of the way, let us now turn to the heart of Field's paper: his attempt to solve the semantic paradoxes of self-reference. In passing, let us note that the paradoxes of self-reference are but one reason for being a dialetheist. There are many other arguments to this effect: arguments concerning Gödel's incompleteness theorem, motion, inconsistent laws and similar things, arguments concerning the limits of thought.[11] Field says (p. 5): 'I don't think there is *any reason whatever* to believe that [there might be some problem solvable better by dialetheist means than by non-dialetheist means]'. Now, it may or may not be the case that the paradoxes of self-reference provide the strongest argument for dialetheism, but these other arguments are not a nothing, and Field has done nothing whatever to show where they fail.

But let us stick with the semantic paradoxes. Field thinks that to accommodate the T-schema one needs a non-classical and detachable conditional (i.e., one that satisfies *modus ponens*). In this we are in full agreement. Such a conditional is given in *In Contradiction.*[12] There is also a wide variety of relevant conditionals that will do the job. (In fact, I now prefer one of these, a depth-relevant logic somewhere in the vicinity of B.[13]) One may even have relevant conditionals in a logic without the LEM, and in which the T-schema

[10] In fn. 19, Field moots the possibility of a family of operators, N_α, such that it gets 'harder and harder' to accept $A \wedge N_\alpha A$ as α increases. Now, even without Field's construction, there is an operator, N, such that accepting A and NA is really hard—irrational, in fact. Let NA be $A \to \bot$. Then one who accepts A and NA is commited to everything. For further discussion of N and its relationship to negation, see Priest (1999), esp. sect. 8.

[11] These are detailed in Priest (1987) and (1995).

[12] Priest (1987), ch. 6. In fact, a couple of options are given there, depending on whether or not one wants to endorse contraposition.

[13] See Priest (2001b), chs. 9, 10. One may show the non-triviality of the T-schema based on such logics by drawing on the work of Brady (1989). See Priest (2002), 8.2.

is demonstrably consistent.[14] We have, then, a multitude of possibilities, and we need to address the question of which is the best. I will turn to this in due course, but first let us see whether Field's construction really does solve the paradoxes.

A standard objection to proposed consistent solutions to the semantic paradoxes is that they all seem vulnerable to 'revenge' paradoxes. There is a certain notion the intelligibility of which the theorist presupposes which, if it is included in the language in question, can be used to refashion the paradox. Hence consistency can be maintained only at the cost of incompleteness—which naturally gives rise to a hierarchy of metalanguages, and so to familiar problems of the same kind.[15] Field claims that his theory is immune from this problem. Is it? The fullest treatment of the point is in Field (2003). In this section and the next I follow his exposition and discussion there. (Though I make no attempt to summarise the technical details of his construction.)

Field's semantics is based on a three-valued logic, where the values are 1, 1/2 and 0. It also employs a (double) transfinite recursion. In terms of this, we may define the ultimate semantic value of a formula, A, $\|A\|$, which is one of these three values. (Ultimate value is value at *acceptable* levels of the hierarchy.) Validity is defined in terms of the preservation of ultimate value 1. For any A, $\|T\langle A\rangle\| = \|A\|$; and given the recursive truth conditions for $\to$, this ensures that the T-scheme holds in the form of a biconditional. Having ultimate semantic value 1 is to be understood as determinate truth—or at least, something like it; the reason for the hedging will be come clear in due course—ultimate value 0 is determinate falsity, and ultimate value 1/2 is indeterminacy. All three notions can, in fact, be defined in terms of determinate truth. $\|A\| = 0$ iff $\|\neg A\| = 1$, and $\|A\| = 1/2$ iff $\|A\| \neq 0$ and $\|A\| \neq 1$. It is, at any rate, the notion of determinate truth and its cognates which threaten revenge paradoxes for Field.

In terms of $\to$, Field shows how to define a family of operators, D^σ (for a certain family of countable ordinals, σ) each of which may be taken to express determinate truth, in some sense. Each predicate in the family applies to some of the sentences that have ultimate value 1, but not others. And each of the standard sentences that are indeterminate, A, such as the

[14]See, e.g., Brady (1983).

[15]See Priest (1987), ch. 1 and Priest (1995), Conclusion.

Liar sentence, the Curry sentence, various extended paradoxes employing the D^σ themselves, and so on, can have their semantic status expressed in terms of some D^σ in the family—in the sense that $\neg D^\sigma A \wedge \neg D^\sigma \neg A$ has ultimate value 1. All these sentences may therefore have their indeterminacy expressed, in some sense. A natural question at this point is whether every indeterminate sentence, B, satisfies $\neg D^\sigma B \wedge \neg D^\sigma \neg B$, for some D^σ, and so can have its status expressed in this way. The answer to this is currently unknown.[16] If this is not the case, then the language is clearly expressively incomplete, since there are indeterminate sentences whose status cannot be expressed. But even if it is, it is not clear that the construction is an advance on the Tarskian one in this respect. In the Tarski hierarchy, each sentence can have its semantic status expressed by some sentence in the hierarchy. Field, it is true, has a single language, not a hierarchy. But this is a superficial difference. One can always think of a hierarchy of languages as a single language.

The expressibility of the status of particular sentences in the language is not the major worry, however. None of the D^σ predicates expresses determinate truth *in general*; and it is this that gives rise to the paradigm revenge problem. Suppose that there were a predicate, D, in the language, such that $D \langle A \rangle$ has ultimate value 1 if A does, and ultimate value 0 otherwise. We then have an extended paradox of the usual kind. By the usual self-referential moves, we could construct the sentence, F, of the form $\neg D \langle F \rangle$. Substituting in the T-schema gives us that $T \langle F \rangle \leftrightarrow \neg D \langle F \rangle$. Now if F is has ultimate value 1, so does $T \langle F \rangle$, and so, therefore, does $\neg D \langle F \rangle$ (the T-schema preserving ultimate truth values); hence F does not have ultimate value 1; conversely, if F does not have ultimate value 1 then $\neg D \langle F \rangle$ does, as, therefore, does F. Deploying the Law of Excluded Middle, we are back with the usual contradiction.

There is a possible move here, with which Field shows some sympathy. Let D express genuine determinate truth; we may add this to the language if necessary. Why must we suppose that it satisfies the LEM? If it does not, the paradox is broken. We still have a revenge problems with us, however. Let F be as before. Suppose that its status is determinate. Then we have Excluded Middle for F, $F \vee \neg F$, and so a contradiction. On pain of contradiction, F cannot be determinate. But its status cannot be expressed by $\neg D \langle F \rangle \wedge$

[16] Field (in correspondence) has conjectured that this is, in fact, the case.

$\neg D\langle\neg F\rangle$. For this, after all, entails $\neg D\langle F\rangle$, that is, F, and so $F \vee \neg F$; in which case we have a contradiction. Consistency is purchased, as ever, at the expense of expressive incompleteness.

4 Enter ZF

But wait. Field's construction is a set-theoretic one, and the language of set-theory can be taken to be a part of Field's object language. The metatheory is therefore expressible in the object language. But this means that ultimate truth value can be defined in the object language—and in such a way that it satisfies the Law of Excluded Middle. And the semantics shows that the theory is consistent. What has gone wrong here?

The answer is that, despite initial appearances, appropriate metatheoretic reasoning *cannot* be performed in the object theory. In one sense, the language may be expressive enough; but in another, and more important, sense, it is not. As Field himself points out, though one can define a predicate 'has ultimate value 1' in the language, this has to be interpreted with respect to the ground model of ZF which kicks off the transfinite construction. And for the usual reason, this model can contain only an initial segment of the ordinals. 'Ultimate value 1', then, means only determinate truth with respect to this initial segment of ordinals, not absolute determinate truth. Field's metatheory cannot be expressed in the object language any more than that of ZF can be expressed in ZF, and for exactly the same reason. If it could be, the theory would be able to establish its own consistency, which is impossible, by Gödel's second incompleteness theorem.[17] Consistency, then, is maintained only by the usual trade-off with expressiveness.

It is worth noting, at this point, that this failure of ability on the part of ZF is effectively its own revenge problem. If ZF could express all that one would expect, it would collapse into inconsistency. Specifically, the inability of ZF to express its own semantic notions (is one of the things that) keeps it consistent. If it were able to show the existence of a universal set, and hence of an interpretation (in the model-theoretic sense) of its own language, inconsistency, in the shape of Gödel's (second) Incompleteness Theorem would

[17]In particular, if $B(x)$ is a proof predicate for the theory—or at least, an appropriate axiomatic fragment of it—one cannot prove $\forall x(B(x) \rightarrow Tx)$, or a consistency proof would be forthcoming.

arise. The fact, then—if it is a fact—that the revenge problem for the theory of truth has turned out to be the same as that for ZF is not reassuring.

In summary, then, though Field may have gone further than anyone else towards the Holy Grail of a consistent semantically closed theory, in the end he fails for the usual reason. The theory, if consistent, is expressively complete.

There is always, of course, the heroic solution: throw away the ladder. We declare the things that cannot be expressed, including the offending determinately-true predicate, to be meaningless. We still, after all, have the set-theoretic construction which can be carried out within ZF, and this at least suffices to show the consistency of the theory of truth relative to ZF. Field shows some sympathy with this possibility too.

The move of declaring the metatheoretic notions meaningless is, of course, open to anyone who wishes to avoid extended paradoxes. And if this is the best Field can do, he is no longer ahead of the field. But the move is one of desperation and would, I think, be somewhat disingenuous. It is clear from the informal way that Field uses the notions of determinacy/indeterminacy in his paper, that these are no mere technical device. Their intuitive sense drives the whole construction. Even Field cannot shake himself free from the meaningfulness of these notions, as the following quotation shows:[18]

> I do not wish to suggest that the notion of having semantic value 1 in the sense defined [i.e., relative to a model] has nothing to do with truth or determinate truth. On the contrary, it serves as a good model of these notions (in an informal sense of model)...

The situation, then, is this. There are notions which, for all the world, appear to us to be intelligible; these cannot, on pain of contradiction, be expressed in the object language. If we declare them meaningless, this is for no reason, in the last resort, other than that they lead to contradiction. As far as solutions to the paradoxes go, the result is, to put it mildly, disappointing. Moreover, and importantly, *ad hoc* manoeuvring of this kind is not required dialethically. A dialetheic (set) theory can contain its own model-theory.[19] We are now coming to the business of a comparison of Field's approach with a

[18]Field (2003), p. 169.

[19]See Priest (200b).

dialetheic approach; before we pursue this matter explicitly, there is another matter to be discussed first.

5 Berry's Paradox

Setting revenge issues aside, there is another, and clear, reason why Field's construction does not provide what is required. Field claims (sect. 5) that his construction can be applied to all the semantic paradoxes, since it generalises in a natural way to the notion of satisfaction, and hence to all semantic notions—including denotation. Now, it is true that the construction can be extended to give a consistent theory of the naive Denotation Schema. However, the denotation paradoxes, such as Berry's, use descriptions essentially. These are not part of the language that Field considers, and their addition blocks Field's construction—at least in any straightforward way. Moreover, the argument for Berry's paradox uses very little logical machinery:– no principles concerning the conditional that are not validated in Field's construction, and no LEM. A proof of this fact is given in Priest (1987), 1.8.[20] Hence, even if Field's construction does solve the Liar paradox, there are equally important semantic paradoxes of self-reference that it does not solve.

In fn. 14 Field replies to this point. He notes that the argument in question uses the least-number principle (LNP):

$\exists x A(x) \to A(\mu x(A(x))$

and argues that this entails the LEM—or at least a restricted version thereof for statements of determinacy, $Det[A] \vee \neg Det[A]$ (I follow his notation here). He therefore rejects this principle. Several points are relevant here.

First, there is something curious about the form of Field's reply. The LNP gives contradiction in the shape of Berry's paradox. If Field were right that it entails $Det[A] \vee \neg Det[A]$, and so gives rise to revenge paradoxes, this would seem to make matters *worse* for him. Field is, in fact, simply rejecting the LNP on the ground that it gives rise to contradiction. The unprincipled nature of this rejection is clear. It is no better than the rejection of the T-schema for no reason other than it gives rise to paradox.

Second, and in any case, the LNP does not entail the LEM. To see this, let '$\mu x A(x)$' denote the least x (in whatever ordering is in question) such

[20] The matter is further discussed in Priest (200a).

that $A(x)$ takes the value 1. For smaller y, $A(y)$ may take the value 0 or some intermediate value. (If there is no such x, we do something else. What, for the case at hand, is irrelevant). This gives us a model of the LNP which does not validate the LEM. Field's argument from the LNP to the LEM uses the principle that:

$$1 = \mu x A(x) \rightarrow \neg A(0)$$

This fails in these semantics (since $A(0)$ may have a value other than 0). Neither is this principle employed in the argument for Berry's paradox.

Third: to see whether the LNP entails the LEM for statements of determinacy, we need to ask how determinacy is to be understood. Perhaps the most natural understanding is to take determinacy to mean having value 1 in Field's own semantics. But if one does this, then $Det[B] \vee \neg Det[B]$ holds anyway, even without the LNP, since set-theoretic statements are two-valued. Rejecting the LNP is therefore beside the point. Alternatively, we may understand it in such a way that it does not automatically satisfy the LEM. Thus, we may understand it as Field's own D operator (D^0). But if we do this, then Field's argument to the effect that the LNP gives $DB \vee \neg DB$ fails. This argument uses the principle that:

$$1 = \mu x A(x) \rightarrow \neg DA(0)$$

But this is not valid. To see this, take $A(1)$ to be any formula such that $DA(1)$, and so $A(1)$, takes the value 1. Take $A(0)$ to be such that both $A(0)$ and $DA(0)$ take the value 1/2 (e.g., the L_1 of Field (2003), p. 159). Then $1 = \mu x A(x)$ holds, but $\neg DA(0)$ fails.

Fourth, one could define the behaviour of the μ-operator differently. We could let '$\mu x A(x)$' denote the least x such that $A(x)$ takes the value 1, and for all $y < x$, $A(y)$ takes the value 0 (or such that $DA(x)$ and, for all $y < x$, $\neg DA(y)$ take the value 1). If one does this, then, of course, the offending principles fall out. But, as I have just shown, there is still an intelligible notion of the least least number operator—intelligible on Field's own semantics—which does not deliver the LEM or a restricted version thereof. And this notion—I emphasise again—is sufficient for the derivation of Berry's paradox.

Fifth, and related, the propreties of the least number operator are a bit of a red-herring anyway. The argument for Berry does not require a least number operator. An indefinite description operator will do just as well. (See Priest (1983).) This demonstrably does not give the LEM, even in the context of intuitionist logic. (See Bell (1993.))

Sixth, and finally, the point remains: descriptions are involved essentially in paradoxes of denotation, and Field has not shown that his construction can handle them in such a way as to give his desired result.

6 Comparisons

Let us now return to the fact that there are many ways of achieving Field's goal: obtaining a conditional that allows acceptance of the T-schema. We need to evaluate which it is more rational to accept. This is to be done (as indicated in section 2) on the ground of which approach is over-all preferable. It is clear that Field's approach is consistent, whilst dialetheism is not. This will be taken by many to be clearly in his favour. In fact, I think that it is not so clear. Both Field and I take classical truth values to be the norm in truth-discourse. We both diagnose the naughty sentences as having some other status. For him, this is a consistent one; for me it is not. But given that we are dealing with an abnormality, consistency does not seem to have a great deal of advantage over inconsistency (especially given that much of the orthodox attitude to inconsistency is a prejudice with no rational ground[21]). However, let us not go into the matter here. Concentrate, instead, on the fact that consistency is not the only relevant consideration.

Another criterion of great importance to the evaluation of rival views is simplicity. A comparison of Field's semantics and those of dialetheism makes it clear where the virtue of simplicity lies. Both approaches deploy what can be thought of as a many-valued logic. But Field's semantics for the conditional involve the complexity of ordinal arithmetic. The construction of Field (2003), employing a three-valued logic with truth conditions that require double transfinite recursion, is perhaps best thought of as a consistency proof for certain principles concerning the conditional and truth. But the construction of Field (200+) is even more complex in some ways, involving as it does a many-valued logic whose values are functions from a set of transfinite ordinals to the three values. Compare this with a dialetheic approach, which deploys nothing more than a certain use of impossible worlds, which there are independent grounds for supposing to be necessary anyway.[22]

[21] This is discussed further in Priest (2001a), sect. 4.

[22] For details of the semantics, see Priest (2001b), ch. 9.

And if I am right about the revenge problem and Berry's paradox, further epicycles must be added to Field's theory to take account of these problems, if this can be done at all.

It also needs to be noted that the conditional that Field comes up with has quite counter-intuitive properties, since it is not a relevant conditional. Thus, for example, it validates the schema $A \vdash (B \rightarrow A)$. Now let A be 'you will not be harmed tomorrow'. This, let us pray, is true. Let B be 'you jump off the top of the Empire State Building tomorrow'. Then from the principle in question we can infer: If you jump off the top of the Empire State Building tomorrow you will not be harmed (then)—which, it would seem, is patently false.

This is one of the milder 'paradoxes of material implication', of course, and there are well know moves that one might make in connection with it (such as an appeal to conversational implicature). But there are many more virulent 'paradoxes', against which such replies are useless.[23] One of my favourites is this. Suppose that we have a light bulb, in series with two switches (currently open), a and b, so that the light will go on if (and only if) both switches are closed. Let L, A, and B be the sentences 'The light goes on', 'Switch a is closed' and 'Switch b is closed', respectively. Then we have: $(A \wedge B) \rightarrow C$. Now, for the material conditional:

$$(A \wedge B) \rightarrow C \vdash ((A \wedge \neg B) \rightarrow C) \vee (\neg A \wedge B) \rightarrow C)$$

If we could apply this, it would follow that there is one of the switches such that if it is closed whilst the other remains open, the light will go on. This is crazy.

Now, the above inference is not valid for Field's conditional, so he might be thought to avoid this problem; but in contexts where we have the LEM for the sentences involved, it does hold. And violations of the LEM arise, for Field, only when the truth predicate, or maybe vague predicates, are involved; but neither of these is the case in the switching example.

Field might say that his account of the conditional was never meant to apply to cases such as this. But this highlights a new consideration. When we select a logical theory as the most preferable, we do not select in isolation: we must bear in mind all the other things that are affected by the choice. Thus, classical logic is simpler than either dialetheic logic or Field's. But, as

[23]See, e.g., Routley *et al.* (1982), p. 6 ff.

Field and I agree, we have to take into account not just the logic, but truth as well. And classical logic plus, say, the Tarski hierarchy of truth (with all the epicycles necessary to avoid its unfortunate consequences) is perhaps more complex that Field's combined deal; it is certainly more complex than a dialetheic approach. We need to take into account a lot more than truth, however. We need to take into account, also, the adequacy of the conditional in other contexts. As we have seen, Field's conditional fares badly in this regard.

Another way in which a dialetheic account is preferable to Field's is in the matter of uniformity. All the standard paradoxes of self-reference are of a kind (inclosure paradoxes). One should therefore expect essentially the same solution for all of them. Same kind of paradox, same kind of solution (the Principle of Uniform Solution).[24] A dialetheic account respects this constraint. Field's account does not. As we have seen, it does not, on its own, resolve the paradoxes of denotation. Some essentially independent factor will have, therefore, to be invoked, violating uniformity.

It is not just the other semantic paradoxes that are inclosure paradoxes. The self-referential paradoxes of set-theory are too.[25] To the extent that Field has a solution to the set-theoretic paradoxes, it is essentially that offered by ZF, which is built into his account. Not only is this solution problematic in a number of ways,[26] it proceeds by denying the existence of various sets,

---

[24]This is defended at much greater length in Priest (1995), 11.5, and the second edition, 17.6.

[25]See Priest (1995), ch. 11.

[26]See Priest (1987), ch. 2 and Priest (1995), ch. 11. Field suggests (section 5) that some of the problems of ZF can be overcome by constructing a theory of properties, on the same lines as his theory of truth. These properties can perform, in a more satisfactory way, the function that proper classes are often invoked to perform. Of course, properties are not extensional, so we lose this much of the theory of proper classes. Field suggest that this is not a problem: 'I doubt that extensionality among proper classes plays much role anyway' (Field (200b), p. 22). In my turn, I doubt that this is true. Here is one reason. *One* of the factors that drives us towards recognising collections other than those that exist within ZF is category theory, where it is natural to consider the category of all sets, or all categories—which are not ZF sets. (See Priest (1987), 2.3.) Now, when reasoning about categories, and in particular when reasoning about their identity, it is standard to deploy extensionality. It is not at all clear that this can be avoided.

a stratagem quite different from that deployed by Field for the semantic paradoxes. Again, we have a violation of the Principle of Uniform Solution.

Let me end this comparison of Field's approach and dialetheism with one further comment. Arguably, a strength of Field's account is that classical logic can be preserved 'where we want it'. In those situations in which the LEM holds, classical logic is forthcoming. Actually, as I have already pointed out, it is not clear that this really is a strength. We have seen that there are good reasons to doubt the adequacy of the classical account of the conditional. But I just point out here that all classical reasoning can be recaptured just as much in dialetheic logic. From a dialetheic perspective, the extensional logic of consistent situations—and classically there are no others—is classical.[27] Another supposed advantage of Field's approach therefore evaporates.

7 Conclusion

We have seen that Field's general worries about dialetheism are groundless, that his approach to the liar is still subject to revenge problems, that it does not even handle all the semantic paradoxes of self-reference, and that, in any case, all the theoretical virtues with the exception of consistency pull towards a dialetheic account of the paradoxes. The 'main case for dialetheism has' not, therefore, 'disappeared' (p. 18). Field's attempt to dispose of dialetheism is markedly less successful than that of Henry II's knights to dispose of Thomas.

> When you come to the point, it does go against the grain to kill an Archbishop, especially when you have been brought up in good Christian traditions.—Third Knight.[28]

---

[27]See Priest (1987), 8.5 and Priest (2002), 7.8. One may still disagree, of course, as to whether conditionals are extensional—especially outside mathematics.

[28]T. S. Elliot, *Murder in the Cathedral,* II: 452-455.

References

[1] B. Armour-Garb and J. C. Beall (2003), 'Should Deflationists be Dialetheists?', *Noûs* 37, 303-24.

[2] J. C. Beall (200a), 'True and False—As If', forthcoming in G. Priest, J. C. Beall and B. Armour-Garb (eds.), *New Essays on Non-Contradiction*, Oxford: Oxford University Press.

[3] J. L. Bell (1993), 'Hilbert's ε-Operator and Classical Logic', *Journal of Philosophical Logic* 22, 1-18.

[4] R. Brady (1983), 'The Simple Consistency of Set Theory Based on the Logic CSQ', *Notre Dame Journal of Formal Logic* 24, 431-9.

[5] R. Brady (1989), 'The Non-Triviality of Dialectical Set-Theory', pp. 437-70 of G. Priest, R. Routley and J. Norman (eds.), *Paraconsistent Logic: Essays on the Inconsistent*, Munich: Philosophia Verlag.

[6] H. Field (2003), 'A Revenge-Immune Solution to the Semantic Paradoxes', *Journal of Philosophical Logic* 32, 139-77.

[7] H. Field (200+), 'Is the Liar Both True and False?, this volume.

[8] H. Field (200a), 'The Semantic Paradoxes and the Paradoxes of Vagueness', forthcoming in J. C. Beall and M. Glanzberg (eds.), *Liars and Heaps*, Oxford: Oxford University Press.

[9] H. Field (200b), 'The Consistency of the Naive(?) Theory of Properties', *Philosophical Quarterly*, forthcoming.

[10] P. Horwich (1990), *Truth*, Oxford: Basil Blackwell.

[11] G. Priest (1983), 'The Logical Paradoxes and the Law of Excluded Middle', *Philosophical Quarterly* 33, 160-5.

[12] G. Priest (1987), *In Contradiction*, Dordrecht: Kluwer Academic Publishers.

[13] G. Priest (1995), *Beyond the Limits of Thought*, Cambridge: Cambridge University Press; second (extended) edition, Oxford: Oxford University Press, 2002.

[14] G. Priest, (1999) 'What Not? A Dialetheic Account of Negation', pp. 101-20 of D. Gabbay and H. Wansing (eds.), *What is Negation?*, Dordrecht: Kluwer Academic Publishers.

[15] G. Priest (2000), 'Truth and Contradiction', *Philosophical Quarterly* 50, 305-19.

[16] G. Priest (2001a), 'Why it's Irrational to Believe in Consistency', pp. 284-97 of B. Brogaard and B. Smith (eds.), *Rationality and Irrationality; Proc. 23rd International Wittgenstein Conference*, Vienna: öbvahpt.

[17] G. Priest (2001b), *Introduction to Non-Classical Logic*, Cambridge: Cambridge University Press.

[18] G. Priest (2002), 'Paraconsistent Logic', pp. 287-303, Vol. 6, of D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, second edition, Dordrecht: Kluwer Academic Publishers.

[19] G. Priest (200a), 'The Paradoxes of Denotation', forthcoming in V. F. Hendricks, S. A. Pedersen and T. Bolander (eds.), *Self-Reference*, Dordrecht: Kluwer Academic Publishers.

[20] G. Priest (200b), 'Paraconsistent Set-Theory and Metatheory', forthcoming in A. Irvine (ed.), *Essays on Set Theory*, Oxford: Basil Blackwell.

[21] R. Routley, *et al.* (1982), *Relevant Logics and Their Rivals, Vol. I*, Atascadero, CA: Ridgeview.