Some Comments and Replies

Graham Priest

April 26, 2019

Contents

1 Introduction: an Orientation

First of all, let me say a warm thank you to all those who have contributed papers to this volume. Many of them are old friends and/or colleagues and/or ex-students and/or coauthors. I feel honoured that they should have considered it worth spending the time and thought required to contribute to the volume; and I thank them warmly for the kind remarks they make about me.[1]

In what follows I shall comment on each of the papers. There is much more to be said about nearly all of them, but given the context, I shall have to restrict myself to what I take to be the central points; and generally speaking, I do not think this is the place to enter into detailed technical issues. As one might expect, I shall have more to say about some of the papers than others. I shall take the papers in alphabetical order of the authors. I have tried to avoid cross-references, though I have used these sometimes where not to do so would have resulted in large chunks of repetition.[2] Occasionally, I have permitted myself a few remarks of a more personal nature.

[1]Many thanks, too, go to all those who sent me comments on earlier drafts of the following sections, which certainly helped improve them. Since I have no better place to say it, let me also say a heartfelt 'thank you' to Çan Baskent and Tom Ferguson for their initiative and all the hard work involved in producing this volume.

[2]I shall frequently refer to what is said, sometimes quoting. At the time of writing, however, I do not have the appropriate page numbers. So I will reference by section numbers.

In what follows, I will have to refer to some of my books a number of times. To avoid prolix referencing, I will refer to them as follows: IC, *In Contradiction* (Priest (1987)); BLoT, *Beyond the Limits of Thought* (Priest (1995a)); TNB, *Towards Non-Being* (Priest (2005a)); DTBL, *Doubt Truth to be a Liar* (Priest (2006a)); INCL, *Introduction to Non-Classical Logic* (Priest (2008)); ONE, *One* (Priest (2014d)). A '2' following the acronym will indicate the second edition. Page references are to second editions, where they exist.

2 Allo and Primiero: Adapting Adaptive Logic

Patrick Allo and Giuseppe Primiero deliver a propositional multiple-conclusion version of Batens' adaptive logic based on the lower limit logic $CLuN$. This is clearly an interesting technical construction. I am content to leave an analysis of it, its strengths and weaknesses, to those who understand adaptive logic better than I do. They locate the construction in the context of classical recapture. Let me say a few things about that.

Paraconsistent logics are normally proper sublogics of classical logic. In particular, the disjunctive syllogism fails: $A, \neg A \vee B \nvdash B$. Yet classical logic seems to work very well in many contexts—for example, classical mathematics, where (one might hope) there are no inconsistencies. It therefore behoves a paraconsistent logician to explain how.[3] To a certain extent, this may be done by an account of the conditional which does not identify it with the material conditional, so that the disjunctive syllogism is not necessary for applying *modus ponens*. However, this will not deal with reasoning where the disjunctive syllogism proper is used.

Adaptive logic is a beautiful and quite general strategy for classical recapture. Not only does it deliver this, but it does so in such a way as to take account of the fact that inconsistencies can have a *local* role in which inferences they invalidate. I learned the idea of adaptive logic from Batens many years ago, and used it to fashion my own adaptive system, LPm.[4]

LPm is a simple construction for showing how the adaptive trick may be turned. Batens' own construction is much more general, and may well have other advantages over LPm.[5] If so, all well and good. We are now discussing

[3]The methodological significance of this was pointed out as long ago as IC, 8.5.
[4]See Priest (1991).
[5]As I point out in Priest (2017), 6.3.

the best way in which the classical recapture can be carried out, not *that* it can be.

Allo and Primiero locate their construction (§6) in the context of Beall's recent thoughts about classical recapture.[6] Here, I agree with them entirely. Beall suggests doing away with the disjunctive syllogism altogether. Where one might normally think to employ it, what one has is simply a choice between the conclusion and a contradiction. General considerations of rationality can determine which of these to accept.[7]

Beall is, of course, right that quite general conditions of rationality, and not just what follows from what, play a role in rational belief. However, adaptive logics show how a default assumption of consistency results in elegant non-monotonic notions of consequence, and how classical recapture can be obtained formally. I can really see no objection to this.[8] In particular, I find Beall's rejection of any notion of non-monotonic reasoning very strange. In real life, most of our reasoning is non-monotonic—or to give it a more traditional name, inductive.[9]

3 Batens: Logic and Metatheory, a Beligian Take

Diderik Batens and I have been friends for most of the years I have worked on paraconsistency. Many times have I visited him and the impressive school of logicians he built up in Gent. And when it was decided to hold the first world congress on paraconsistency, Gent was the natural place for it, and Diderik organised the historic conference. Of course, we have discussed paraconsistency and related issues for all these years. I have learned much from these discussions,[10] and I'm delighted to be able to carry on our conversation here. As he says, we tend to come at things from rather different perspectives; and that difference is certainly too big an issue to take on here. We also disagree about some more particular things, though I often think that there is much more agreement between our views than he does. Often, it seems to me, it is just a matter of reorientation. We may disagree about that too!

[6]See Beall (2011), (2012).

[7]The strategy goes back to IC, 8.5.

[8]See Priest (2017), §6.2.

[9]See Priest (2012a).

[10]For some of the discussion, see Priest (2014a).

Anyway, Batens' paper here is rich with ideas and arguments. I can take up only a few of the issues he raises. I will concentrate on what I take to be the two main topics: logical pluralism and paraconsistent metatheory.

3.1 Logical Pluralism and Related Issues

As far as deductive logic goes, I am a logical monist.[11] That means that given some inference there is, in principle, one correct answer to the question, when suitably understood, of whether or not it is valid. Call this *the Question*. I am not against logical pluralism (the denial of monism); if it turns out to be the case, so be it. I don't think that this affects the nature of, or arguments for, dialetheism at all. It is just that, so far as I can see, an adequate case for pluralism has not been made; and methodologically, monism is the simpler view.[12]

Of course, I am well aware that there are many pure logics (classical, intuitionist, paraconsistent, etc). And as pieces of pure mathematics, these are all equally good. Moreover, suitably understood, they can all be thought of as providing answers to the Question. That, of course, does not imply that there is more than one correct answer. Neither do I think that what theory logicians take to be right cannot change over time; of course it can: it has. Nor do I wish to suggest that I or anybody else has the logic which does always answer the Question correctly—though of course, we have reasons to suppose that some theories are better than others. In all these ways, logic is much the same as physics.[13]

What does actually follow is that there is a unique deductive *logica ens*—something that may be distinct from both our theories and our practices. Batens says (§2) that he does not understand this. I find this rather surprising. It is, after all, what logicians have been trying to characterise since Aristotle and the Stoics. Perhaps it is impossible to *achieve* what they have been trying to do; however, *what* they are trying to achieve sees perfectly intelligible.

Naturally, inferences are expressed in language, and, as Batens notes, the

[11]See Priest (2001).

[12]I guess I've always been disposed to accept monism, just because in the early days of paraconsistency, many paraconsistent logicians, such as da Costa, endorsed pluralism as a way of attempting to legitimate what they were doing. This always struck me as a failure of nerve.

[13]For a more general discussion, see Priest (2014b).

languages we actually speak (as opposed to some formal language), are highly idiomatic and ambiguous. It may well be that the correct response to the Question is, 'It depends what you mean'. And this meaning may need to be clarified before a sensible answer is possible. A standard logicians' assumption is that this has already been done before the Question is addressed.[14] But once it is done, one may fairly expect a straight answer to the Question. To say that the inference is valid in classical logic, but not intuitionist logic, would reasonably be seen as avoiding the question.[15]

Of course, again as Batens points out, the meanings of our words can change over time, as, then, may answers to the question 'What do you mean?' That does not imply that the answer to the Question changes. Nor does it imply that there is no such thing as determinate meaning—Derrida notwithstanding. It just means that the results of clarification may change with time. To put it in very traditional terms, an answer to the Question requires us to determine what propositions ares being expressed by the sentences involved. And of course, in the process of change, new concepts may be introduced. A new word is coined, or the meaning of an old word is revised, to express a concept not hitherto expressible. This is a way in which the *logica ens* may change. It can be augmented, simply because new propositions are coming into play.[16]

Batens ties his view to a patchwork view of knowledge (perhaps better, rational belief). Our knowledge about the world is no unified whole—and maybe never will be; it is a patchwork of sometimes inconsistent views. With this, I am completely in agreement. This is one reason why Bryson Brown and I invented the methodology of Chunk and Permeate.[17] As far as I can see, all this is perfectly compatible with what I have said above, which does not mention epistemology at all. Of course, as our theories change, new words may be coined, or old words may come to have new meanings. But as I have already said, this is quite compatible with logical monism.

[14] It is no accident that much philosophical discussion concerns clarification of meanings. It is a sensible thing to do before any kind of debate—and not just in philosophy.

[15] Imagine that a logician is called as an expert witness because one of the arguments used by the defence is particular complicated. The judge asks the logician whether the conclusion of the argument actually follows.

[16] Whether it can change in other ways may depend on what, exactly it is that determines what follows from what. (See Priest (2014b), §4.1.) If one takes validity to be a relation between abstract entities (such as propositions or mathematical structures of a certain kind—as in Priest (1999)), then presumably not.

[17] Brown and Priest (2004).

Which brings me to adaptive logic, a subject for which I have enormous admiration. Adaptive logic is a species of default reasoning, or, to give it a more traditional name, inductive reasoning. This is non-monontonic, in the sense that from certain information we may correctly infer a conclusion because, in some sense, that would be the case in a normal situation. The conclusion can be rescinded, though, given further information, to the effect that we are not in a normal situation. The hackneyed example: Tweety is a bird; so Tweety flies—which is fine until and unless we learn that Tweety is a penguin.

Developments in non-monotonic logic are a great development in 20th century logic. Such logics can be handled semantically by having a normality order on interpretations—the valid inferences being the ones where the conclusion holds in all models of the premises which are as normal as possible, given those premises.[18]

Non-monotonic logic is not a rival to a deductive logic. Such a logic standardly has a deductive logic as its basis. (Batens calls this the *lower limit logic*.) A non-monotonic logic is built atop of this, and *extends* the valid inferences, by adding a normality ordering. The most usual non-monotonic systems take classical logic as the lower limit logic, and use various empirical default assumptions to generate the notion of normality involved.

One of Batens' great achievements was to take, instated of classical logic as the lower limit logic, a weaker (often paraconsistent) logic, to which might be added, not empirical default assumptions, but default assumptions about what one might call logical normality. Prime amongst such assumptions is one of consistency. Adaptive logic is, hence, a very general approach to non-monotonicity; and over the years Batens and his school have clearly shown how such logics can be used for many kinds of default assumptions, including such things as monosemy—as mentioned in his paper.[19]

I have described the basis of adaptive logics in some detail so that, as I hope becomes clear, there is nothing in such a project with which I feel the need to disagree. On the contrary, I have used it in my own limited way, to construct the adaptive logic LPm. (See §2 above.) In particular, there is nothing in this project which in any way challenges monism about deductive logic—any more than inductive logic challenges deductive logic; and there is certainly nothing in this view which is anti-formal-logic: adap-

[18] See Priest (1999).

[19] For a deductive logic which handles ambiguity of denotation, see Priest (1995b).

tive logics *are* formal logics; nor need there be any tension between adaptive logics and paraconsistent logics. Adaptive logics often *are* paraconsistent (in that, according to such logics, contradictory premises do not imply everything). Whether adopting an adaptive logic delivers a defensible solution to the paradoxes of self-reference is far too large a topic to take on here.

3.2 Metatheory

Let me now turn to the question of the semantics of deductive paraconsistent logic, and particularly the semantics of LP. This is the topic to which Batens turns in the second half of his paper.

To address matters here, there is a crucial prior issue. Batens is concerned with the model-theoretic definition of validity. This is formulated in set-theoretic terms, so one has to determine what account of set theory to endorse. A dialetheic account of the set-theoretic paradoxes accepts the naive comprehension schema:

- $\exists x \forall y (y \in x \text{ iff } A)$

where A can is arbitrary.[20] The most crucial question is how to interpret the 'iff'. There are currently two possible views on the table.[21]

One option is to take the underlying logic of the theory to be a relevant logic, and take the 'iff' to be a relevant biconditional. This approach was pioneered by Routley.[22] I have certainly contributed to this project. But undoubtedly the idea has been developed in its most sophisticated form by Weber.[23] Quite how much metatheory one can construct in this approach is still not known. However, for such an approach, the consequences spelled out by Batens in the first part of §4 do, indeed, seem to follow (if the relevant biconditional contraposes). Indeed, related issues were pointed out by Weber.[24] How damaging these consequences are would require a substantial

[20] One normally insists that x not occur free in A, but in a relevant logic this actually implies the more general condition. The argument for this based on an underlying substructural logic can be found in Cantini (2003), Theorem 3.20. The same argument applies to a relevant logic.

[21] Maybe more if one goes substructural.

[22] Routley (1977).

[23] Weber (2010), (2012).

[24] Weber (2016a). However, Weber endorses the axiom that truth and falsity in an interpretation are exclusive, which seems to exacerbate the matter.

discussion. However, I forego this here, because I am inclined to a different approach to naive set theory.

This is to take the underlying logic of the theory to be LP, and to take the 'iff' to be its material biconditional.[25] The conditional of LP does not detach. Hence, there is no hope of trying to prove set-theoretic theorems in this theory. One must sail on a different tack. This is itself model-theoretic.

One can prove that there are models of this theory which verify, not only the naive comprehension schema, but also *all* the theorems of ZFC. If one assumes that the universe(s) of set theory is (are) represented by such a model (models),[26] then one can simply take over all the result of ZFC, including standard metatheory—including that of LP.

And for such an approach, the results about validity set out by Batens do not go through. To establish the results about the relation R he uses would appear to use invalid arguments. Thus, to show the existence of an inconsistent R, one would have to reason as follows. Suppose that k is the Russell set, $\{x : x \notin x\}$, and let K be $k \in k$. Then $K \wedge \neg K$. By naive comprehension, we may define a propositional interpretation, R, such that:

- $\langle x, y\rangle \in R \equiv (x \in \mathbb{P} \wedge (y = 0 \vee y = 1) \wedge K)$

where $\mathbb{P}$ is the set of propositional parameters. One cannot, however, infer that $R(p,1)$, $R(p,0)$, $\neg R(p,1)$, or $\neg R(p,0)$, since the material biconditional does not detach.

This does not mean that logical consequence is a consistent notion, however. The argument for this was given by Young,[27] and is gestured at by Batens in the final pages of the section in question. The argument requires that the set theory should be able to establish the existence of a standard model; that is, a model $\mathfrak{M}$, such that for any formula A:[28]

- $[\mathfrak{M} \Vdash A] \equiv A$

For then, by self-referential techniques, one can find a sentence, D_0, such that:

[25] This approach is explained in the second edition of IC2, ch. 18, and at greater length, in Priest (2017), §§10-12.

[26] Or at least, such models of a relatively low degree of inconsistency—to rule out, for example, the trivial model.

[27] Young (2005).

[28] Of course, this cannot be done in ZFC—assuming it to be consistent. But one can show that there are models of naive set theory and ZFC in which this is the case. (See Priest (2017), §11.)

- $\mathfrak{M} \Vdash D_0 \equiv \mathfrak{M} \nVdash D_0$

It follows in LP that $\mathfrak{M} \Vdash D_0 \wedge \mathfrak{M} \nVdash D_0$, and so $\neg(\mathfrak{M} \Vdash D_0 \supset \mathfrak{M} \Vdash D_0)$. Given that the validity of an inference from A to B is defined as $\forall x(x \Vdash A \supset x \Vdash B)$ (sticking to the one-premise case, for simplicity), it follows that $p \nvDash p$ (though $p \vDash p$ as well), since the inference has a substitution instance which is a countermodel. The obvious generalisation of this argument to other forms of inference does not go through. Whether the inconsistency of the validity relation spreads further still requires investigation.

Batens' objection at this point is simply that the validity relation is inconsistent, contrary to my view.[29] However, as far as I can recall, I have never suggested that it was. Indeed, given that we have a standard model—and so, in effect, a truth predicate—and techniques of self-reference, this is exactly what should be expected. Neither does this seem to me to be particularly problematic. After all, the inference in question is still valid.[30]

Finally, Batens points out that worries about the inconsistency of logical consequence might spill over into worries about non-triviality—and specifically that one might be able to prove (rather trivially!) that every theory is non-trivial; that is, for any T, there is some A such that $T \nvdash A$. Given that we are allowed to assume standard results about soundness and completeness, this means that for some $\mathfrak{A}$, $\mathfrak{A} \Vdash T \wedge \mathfrak{A} \nVdash A$. Now, given Young's argument, it is true that if the language of T is that of set theory, then $\mathfrak{M} \nVdash D_0$; but in general it will not be the case that $\mathfrak{M} \Vdash T$, since $\mathfrak{M}$ is a very particular model.[31]

4 Berto: Impossible Conceiving

Franz Berto and I have worked together on the version of noneism that I and I favour for some years now, and it has been a very fruitful collaboration. He has now started to think more about imagination, and I'm very happy to

[29] He also asks how results about the semantics can have been established in ZF, if they are inconsistent. This they cannot, since they use unrestricted comprehension, and so go beyond ZF.

[30] For a further discussion of the above these matters, see Priest (201+a).

[31] Moreover, even if such "cheap" proofs of non-triviality *were* available, this does not undercut the value of more substantial proofs. For such proofs normally establish not only non-triviality, but also limitations on the class of sentences in the language which T can show to be contradictory.

able to push that project forward. What is at issue in his paper is whether one can imagine the impossible. I certainly think you can; so does he. He thinks that things might be a bit more complicated than I have suggested, though.

I think that conceiving a state of affairs and imagining it are much the same thing.[32] To conceive of something is simply to bring a representation of that state before the mind. (As Berto notes, this does not imply that one can conceive of every impossibility, or even every possibility. There may be some states of affairs—possible and impossible—that transcend anything I can represent to myself, perhaps because they are too complex.[33]) Berto thinks that imagining is not quite the same as conceiving: it is a special sort of conceiving: conceiving-as-imagining.

In conceiving as imagining, the kind of representation one brings before the mind is, in some sense, a pictorial image. A paradigm example of this is when I imagine Socrates drinking hemlock. When I do this, I have a visual image of an old guy with a beard and a snub nose sitting on a couch with a cup in his hand. The image involved in conceiving as imagining does not have to be visual, however. It could be an auditory image—for example, of an orchestra playing Beethoven's *Ode to Joy*—or a kinesthetic image—for example, of my performing a karate kata. One might say that in such imagining, one runs a sensory system in the brain, "but offline". By contrast, in conceiving-as-not-imagining the representation is a linguistic one. Thus, I might conceive its being the case that intuitionistic logic is correct. No sensory image is involved. The representations in this case are linguistic statements such as the sentence 'Excluded Middle is not valid', and so on.

Now I am quite happy to say that conceiving of the latter kind *is* imagining. This sort of imagining is exactly what I do when I imagine intuitionist logic to be correct. But I am happy to agree that the representations involved in imagining can be linguistic or sensory—though I think that the distinction may be vague. Diagrams and maps, for example, seem to have elements which are both pictorial and linguistic. (Thus, when I imagine that intuitionist logic is correct, I might imagine a diagram of Kripke countermodel to Excluded Middle. There is certainly a visual image involved; but the failure of Excluded Middle is not something that one can literally see.)

[32]At least in the relevant sense of 'imagine'. There is a sense of the word in which you imagining something implies that you are not certain of it. That is not the sense in question here.

[33]On all these thing, see TNB2, 9.6.

Set all this aside, however. Berto thinks that conceiving-as-imagining puts up stiffer resistance to the thought that one can imagine the impossible. Thus, suppose that it is a necessary truth that water is H_2O, and that I imagine that it is not. I might have a mental image of this wet stuff, such that once one zooms in, the molecules have some other constitution. One might simply aver, as some have, that it was not *water* that I imagined, but some other wet stuff. Berto's reply is that even in pictorial representation, there is more to matters than the phenomenology. Thus, if I imagine that Hillary Clinton won the 2016 US election, I have a visual image of her celebrating, surround by streamers, etc. But how do I know that it is her, and not just someone who looks like her? Well, *ex hypothesi*, it is *her* I am imagining, not a *doppelgänger*. Similarly (as Berto notes), when I imagine that water is not H_2O, it is *water* that I am imagining, and not some *doppelgänger*.[34] As Berto puts it, even in the pictorial case, it is not like looking at the situation through a telescope; there is an element of fiat about who or what it is that are the objects in the picture.

I think that Berto is quite right about this. However, it seems to me that even without this element of fiat, one can have this kind of imagination of the impossible. We are all familiar with visual illusions. One of the most interesting for present purposes is the waterfall effect. A perceptual—usually visual—system is conditioned by constant motion in one direction. Once this is taken away, one gets a negative after-image, of whatever one is looking at moving in the other direction. But if one focuses one's attention at a point in the visual field it appears to be moving in that direction *and* be stationary as well.[35] That is exactly how subjects of the illusion describe what they see, and you will too if you undertake the experiment. Now something's being stationary and being in motion (in the sense in question) is an impossibility. Yet I can perfectly well picture what it is like for something to be stationary and in motion, because I have seen what it is like: I have done the experiment. Yet, no element of fiat is involved in this: that's just how it looks—as through a telescope.

For good measure, I note that there are auditory phenomena of the same kind. One can generate the sound of a note that appears to be perpetually rising, but stays constant.[36] That is of course, impossible. Now, at the

[34]See TNB2, p. 195.

[35]For discussion and references, see DTBL, §3.3.

[36]See Shepard (1964), Tenney (1969).

end of the Beatles' 'Day in the Life' the orchestra plays a note that actually does gradually ascend. But I can imagine listening to the Beatles 'Day in the Life' when the note which the orchestra plays at the end perpetually rises but remains constant. There is an element of fiat involved in this. *Ex hypothesis*, it is the *Beatles*' 'Day in the Life' that is the subject of my imagining. I am not imagining a situation in which the Rolling Stones composed 'Day in the Life'. But there is no fiat involved in the notes' doing what it does. I know exactly what it sounds like.

A final comment on granularity. I agree with Berto that representations will have a granularity. Some will be more detailed than others—though any representation is likely to be partial, and so silent on some details. However, I do think that I can 'imagine building step by step a perfectly valid proof' of Goldbach's Conjecture (§7). This is all in my imagination. So the validity itself may be imaginary. What I cannot do is imagine building a step by step perfectly valid proof, such that it is actually valid. Were I able to do this, then I would indeed have solved the Goldbach problem.[37] Imagining a state of affairs, as Berto notes, cannot (perhaps sadly) make it obtain.

5 Brady: True or False (Only) Strikes Back

Ross Brady has proved many impressive results in relevant/paraconsistent logic. I still regard his proof of the non-triviality of naive set theory as one of the milestones in the development of the subject.[38] Perhaps one of the main things he and I have argued about over the years is whether the paradoxes of self-reference are best handled via truth value gaps or truth value gluts.[39] His essay in the collection puts the issue in the more general context of how many semantic values there actually are.

First, let me get some distracting matters out of the way. The paper he refers to as the one I gave in Istanbul in 2015 is, I think, Priest (2015). In that, I did not defend the idea that logic should have four—or five—values. I argued that the best way to make formal sense of the Buddhist *catuṣkoṭi* is to do so in a 4-valued logic, and specifically, *FDE*. The paper also discusses

[37] Van Inwagen's objection fails none the less because, *ex hypothesi*, it is [building a valid detailed proof of the Conjecture] which I am imagining.

[38] This appeared as Brady (1989), but the result is much earlier. I remember that we had a whole mini-conference around it in Canberra in 1979.

[39] See IC, ch. 1.

adding a fifth value, e, ineffability—though, as it briefly indicates, if one does so, one has to think of the bearers of values as states of affairs, not as sentences.[40] Given that, we are not considering *semantic* values. In particular, this has nothing to do with the fifth value as a value for linguistic nonsense of any kind, for which I hold no brief.[41] Indeed, I have never argued that we need more than four semantic values. I quite agree with Brady that we do not.

Next, Brady argues that validity is about deducibility. For me, validity is about truth- (or, more precisely, satisfaction-) preservation. This is too big an issue to take on here.[42] I note briefly only two things. First, Brady's claim (§2) that truth-preservation is about propositions, which are, by definition, neither true or false, and not both, is about as tendentious a definition as I think one might find (whether or not it is standard). Propositions (as opposed to questions and commands) are the *kind* of thing that can be true or false; but this says nothing about whether any achieve neither or both of these statuses.) Secondly, on a truth-theoretic semantics there is absolutely no problem about disjunction or the particular quantifier. $A \vee B$ is true iff A and B is true. Even if we know that $A \vee B$ is true, it does not follow that we should know which of these it is, much less that it be proved.[43]

Anyway, set these matters aside. Given any theory, for any A, there are, as Brady says, four possibilities:

1. A is provable and $\neg A$ is not
2. $\neg A$ is provable and A is not
3. Neither is provable
4. Both are provable

And there are theories which deliver each of these possibilities for various As. This brings me to the most significant point of disagreement concerning

[40]The matter is discussed at much greater length in Priest (2018a), ch. 5.

[41]Indeed, I am happy to take (atomic) such sentences, if they occur in the language at all, simply to be false. See IC 4.7.

[42]I have discussed the matter in DTBL, ch. 11.

[43]If anything, it is the proof-theoretic semantics which is problematic. For given such an account of validity (of the kind one finds in, for example, intuitionist logic), $A \vee B$ is provable iff A is provable or B is provable; and this may well not be the case in the relevant logics of the kind that Brady favours (as he, himself, points out).

what Brady says in his essay. He wishes to eliminate possibility 4. I certainly do not.

He says (§1) that:

> the case for dropping the contradictory value [4 ...] will depend on ideal formal systems that represent conceivable concepts and would involve reconceptualising any concept or concepts that lead to contradiction.

In other words, the value is to be eliminated in an ideal situation. I suppose the most obvious thing to say is that even if the value is not present in an ideal situation, we are not normally in one, and hence we have such values. I will return to this matter in due course; but the crucial question is why one should suppose that eliminating case 4 is an ideal, that is, something for which one should strive.

The main considerations for this are marshalled in §3. The first is that Hilbert intended that consistency was what formal systems were meant to achieve. Well, yes. That was his agenda in the philosophy of mathematics; but that was only ever one such agenda; it is now gone; and in any case, Hilbert knew nothing about paraconsistent logics.

The next observation is that ideal logical systems are meant to capture logical notions such as conjunction and negation with conceptual clarity. We may certainly agree with this; but there is nothing unclear about a dialetheic theory of negation, or about something and its negation both being provable.[44]

Next, we are told that when we have a logical system in which something and its negation are both provable:

> in order to avoid inconsistency, people would be inclined to re-examine the concepts to see if such a system can be made consistent by fixing up the axiomatization. It would be thought that a conceptual clash between concepts would have taken place or a particular concept would have been over-determined.

[44] In §4 Brady says that Boolean negation is the 'intended negation'. Intended by whom? Certainly not be me. Boolean negation is a theory of how negation behaves—just a false one. (See DTBL, chs. 4, 5.) Brady tells me in discussion that what he meant was that Boolean negation has 'conceptual completeness', in the sense that the conditions under which something is true also determine the conditions under which its negation is true. (They are simply the complement.) This is certainly so; but that negation behaves like this is simply part of the same false theory.

I am not sure what people Brady has in mind here—wise people, dogmatic people, 20th century people? Certainly not everyone has this tendency in all cases. Even scientists feel no need to render their concepts consistent if those concepts do their job properly. The 17th and 18th century infinitesimal calculus was known to operate on the basis of an inconsistent notion of infinitesimal; but no one at the time felt the need to revise this.[45]

But in any case, since this is about what should be the case ideally, the point is not about what people *do* do, but about what they *ought* to do. Why should such revisions be made? We are told that such a contradiction shows that a concept has been over-determined—implying, I presume, that it has been incorrectly charactised. But why should it not be in the very nature of a concept to be contradictory? *Prima facie*, the liar paradox shows exactly that the notion of truth is over-determined in this sense.

Perhaps, for some inconsistent concepts, there would be good reason to revise them. Thus, consider the concept:

- *x has priority of way at a road junction*

If this were inconsistent, in such a way that different drivers both had priority, it would be sensible to revise it. The point of traffic laws is a practical one; and the law embedded in this concept would certainly have impractical consequences. But such considerations do not generalise. Grant that the concept of truth is inconsistent. Why should we revise it? The concept of truth serves us perfectly well as it is. It does not cause death on the roads.[46]

Later on in the same section, Brady appears to offer another argument. It is better if our concepts are conceivable, and inconsistent concepts—such as *round square*—are not conceivable. So inconsistent concepts should be revised. However, inconsistent concepts may well be quite conceivable. Something round and square may not be *visualisable*. But such is true with many quite consistent concepts, such as that of being a chiliagon. Indeed, the concept *round square* is quite conceivable. We conceive it in saying that anything that satisfied it would have inconsistent properties—we must understand it to know that this is so; and if this is not to conceive it, I have no idea what conception is. Similarly, the notion of truth, if it be inconsistent, is still quite conceivable.

[45]It is true that the notion was abolished in the 19th century with the consistent notion of a limit; but this was done for quite different reasons. See Lakatos (1978).

[46]See IC, 13.6.

In §5 of his paper Brady adduces one further argument for the special case of contradictions in a metatheory.[47] There can be no contradictions to the effect that something is both provable and not provable. First, I note that many statements in a metatheory, as the term is usually used, are not about provability at all. Hence, even if the point about provability were true, it would not rule out contradictions in a metatheory in general. However, I find Brady's argument for his conclusion unpersuasive. If I understand it right, it goes like this. To show that something is provable one gives the proof. One cannot show that something is unprovable in this way. To do so, one needs some quite independent kind of proof procedure (algebraic, model theoretic, etc.), one that is 'outside of the recursive proof process'. Hence 'non-proof cannot overlap with proof'. It seems to me that this does not follow at all. Indeed, the fact that proof and non-proof are to be established by independent mechanisms opens up the very possibility that they may give conflicting results.[48]

So much for case 4 of our four cases. Let me say a final world about what Brady says about case 3: neither true nor false. In §4 he argues that in ideal cases—albeit ones that may not be realisable—there should be nothing in this category either.

Now, first, I entirely agree with him—and not just in an ideal case, but in the actual case! Though one may make a case for truth value gaps, I have never accepted this.[49] Next, I must say that this part of Brady's paper particularly surprised me. For years, Brady has been arguing that the correct solution to the paradoxes of self-reference is to recognise the paradoxical sentences as neither true nor false.[50] Given what he is now arguing, if it be the case that a dialetheic solution to the paradoxes of self-reference does not hold in the ideal case, it is also the case that neither does his! Thirdly, given his background assumptions, to realise the ideal situation for case 3, the theory in question has to be decidable. For many concepts, such as validity in first-order logic, that is entirely impossible. Given so, how can one be sure

[47]In §6 he also avers that 'metatheory is currently assumed to be two-valued'. Well, it is not so assumed by intuitionists or by those who hold that metatheory can be carried out in a paraconsistent metalanguage. See, e.g., Dummett (1977), ch. 5, and IC2, ch. 18. But in any case, have we never heard of current assumptions being wrong?

[48]A quite different argument to the effect that something cannot be provable and not provable is given by Shapiro (2002). I have replied to this in IC2, 17.8.

[49]See IC, ch. 4.

[50]A view which, he tells me, he still endorses.

that the same is not true of case 4? Maybe in many cases it is impossible to eliminate contradiction.[51] Indeed, this would appear to be exactly what extended paradoxes of self-reference have taught us concerning truth.[52]

6 Carnielli and Rodrigues: Expressing Consistency (Consistently)

In their paper, Walter Carnielli and Abilio Rodrigues present natural deduction systems for the logics they call BLE and LET_J. Essentially, the first is Nelson's Logic N_4, and the second augments this with a classicality operator to give an LFI. They claim that these deliver a proof theoretic account of meaning, and can be motivated in terms of evidence-preservation. Much of the material, together with more technical details, appear in Carnielli and Rodrigues (2017). I shall leave to those who are concerned with proof-theoretic semantics the question of whether these systems are adequate for that purpose. Here, I will just comment on a few other philosophical issues they raise, end especially those connected with dialetheism.

First a preliminary issue. Carnielli and Rodrigues (hereafter, C&R) note that the use of a paraconsistent logic (such as N_4) does not commit one to dialetheism. Indeed not. This is something that I have pointed out many times.[53] Nor does using LP imply a commitment to dialetheism either. One may hold that the actual world is consistent, and that the interpretations in which contradictions hold represent impossible situations. Indeed, there are interpretations of LP which dispense with the middle "contradictory" value altogether.[54] The interpretation of paraconsistent logics where the semantic values are given informational interpretations ('told true' and 'told false') are also well known.[55] Having said that, the mere fact that there are interpretations of a paraconsistent logic, such as theirs, which does not endorse dialetheism is not an argument against it. There may be interpretations

[51] Further on that matter, see Priest (2014a).

[52] In §6 of his paper Brady briefly addresses the matter of extended paradoxes. His position, if I understand it correctly, is to resort to the Tarskian object/metalanguage distinction. But this move, apart from being problematic in its own right, as shown by Kripke (1975), does not eliminate extended paradoxes. See IC, 1.5.

[53] See, e.g., Priest (2002), §2.2.

[54] See Brown (1999).

[55] See, e.g, Belnap (1977).

of it which do; or, alternatively, if dialetheism is correct, this may simply be the wrong logical system. C&R do not engage with the arguments for dialetheism at all.

Next, they attribute to me views about dialetheism which I do not hold.[56] A dialetheia is a pair of sentences (or their conjunction, if you like) of the form A and $\neg A$, such that both are true. If we take 'false' to mean having a true negation, this means that A is both true and false. Dialetheism is the claim that some As are dialetheias.[57] The definition is not committed to any particular view of truth. That dialetheism presupposes no particular theory of truth is spelled out at length in DTBL, ch. 2. There, I point out (among other things) that one can have a verificationist view of truth, that is, one that is "epistemically constrained"—where truth is warranted assertibility—of the kind espoused by anti-realists.[58]

It is therefore mistaken to claim that 'the dialetheist claims that some contradictions are ontological in the sense that they are due to some "inner contradictory essence of reality"'.[59] Indeed, as IC2, 20.6 points out, it is not even clear that the claim that there are contradictions in reality makes sense.[60] For it to do so, one has to endorse some kind of correspondence theory of truth, holding reality to comprise facts or fact-like entities.[61] I have never endorsed such a view. Indeed, the only theory or truth I have ever advocated (IC, 4.5—the "teleological theory of truth"), is anything but such a realist theory.[62] So let me say it one more time: it is not clear that a philosophically substantial claim to the effect that there are contradictions in reality makes sense; and even if it does, I am not committed to it.

[56]In fairness to them, these are views I hear not infrequently. I have no idea why this is so, since the views are without textual support; indeed, they are against textual support.

[57]This is how matters are defined in IC, p. 4.

[58]On anti-realism and truth, see Glanzberg (2013), §4.

[59]Carnielli and Rodrigues (2017), p. 1. I have no idea from where they are drawing the quotation. This is certainly not something I would say.

[60]Except in the entirely banal one that for a sentence to be true it requires the cooperation both of words and of the world. Thus, 'Brisbane is in Queensland' is true because of the meanings of 'Brisbane' and 'Queensland', and of Australian geography.

[61]Matters are made worse by the fact that C&R mis-state this view. They say (§8): 'let us recall that a true contradiction would be made true by an object a and a property P such that a does and does not have the property P'. No. This is to import classical truth conditions where they are obviously not apt. a should be in the extension of P and the anti-extension of P, which is quite different from not being in the extension of P.

[62]And I point out that the account is compatible with different kinds of sentences having different kinds of truth-makers (IC, p. 57).

I now turn to matters concerning meaning. Broadly speaking, there are two contemporary approaches to sentence-meaning. One is that the meaning of a sentence is determined by its truth conditions.[63] The other is that it is determined by rules of use, and especially proof. The first line was endorsed, famously, by Frege, and later taken up by Davidson. The second was pioneered by Gentzen. I have endorsed a version of the first view.[64] C&R endorse a version of the second. (Though what view of truth they endorse is not clear to me; nor is the connection they envisage between truth and meaning. It is certainly not of the intuitionist kind, since they explicitly distinguish between warranted assertibility and truth.) That particular difference is too big an issue to take on here, though I note that nothing about dialetheism presupposes a truth-conditional account of meaning either. Indeed, given a proof-theoretic account of meaning, if our rules for the use of words (or our rules of proof)—maybe those concerning the T-Schema—are such as to establish A and $\neg A$ then dialetheism holds, since these things are true in virtue of meaning, whatever theory of truth one endorses. C&R's claim that their logic does not support dialetheism (§8) because one cannot have (non-trivially) $\neg A \wedge \circ A$ and $\neg A \wedge \circ A$. But one can have A and $\neg A$, so this observation seems to miss the point.[65]

Let me make one further remark about truth-theoretic views of meaning.[66] C&R, as do many people, appear to conflate truth *simpliciter* (*ts*, from now on) with truth-in-an-interpretation (*tii*, from now on). (The fact that Tarksi wrote seminal papers on both tends to abet this confusion.) *ts* is a non-relational notion that satisfies the T-Schema (or at least some close cousin.) *tii* is a set theoretic relation, and it has one major function: to deliver a model-theoretic notion of validity. Of course, one might hope that there is one interpretation such that truth in it coincides with *ts*, but this may not be the case. Witness the fact that the set theory ZFC can give a definition of validity for its own language, but on pain of inconsistency it cannot establish the existence of such an interpretation, by the Tarski indefinability proof.

[63] Or perhaps truth-in-a-possible-world conditions, if we are dealing with modal notions.

[64] See IC, 9.4.

[65] Unless they are identifying truth/falsity with classical truth/falsity. That would clearly be entirely question begging in the context of any argument against dialetheism.

[66] Also, a minor comment. C&R claim (§5) that classical logic cannot be given a proof theoretic semantics. This is false. See Read (2000). One merely has to find an appropriate notion of proof-theoretic harmony.

Now, the notion of truth (and coordinately, of reference) that is at issue in a truth-conditional account of meaning is *ts*, not *tii*. C&R criticise the idea that model theory has anything to do with meaning. Quite rightly. But it was never supposed to do so. The appropriate notion of truth for this job is *ts*.

A final comment on what C&R say about evidence. They argue that the notion of validity in their preferred logic can be motivated in terms of the preservation of evidential support, by analogy with the way that validity in intuitionist logic can be motivated in terms of preserving provability. There are reasons to suppose that this motivation does not succeed. For evidence, unlike proof, is defeasible. A can provide evidence for C, but B can override this. (So 'Tweety is a bird' is evidence that Tweety flies. But 'Tweety is a bird and Tweedy is a penguin' is not evidence that Tweety flies—quite the opposite.) In other words, evidence is non-monotonic. However, in C&R's logic, if $A \vdash C$ then $A \wedge B \vdash C$. For good measure, the logic validates adjunction $A, B \vdash A \wedge B$; but there appear to be perfectly good situations where we can have evidence for A and evidence for $\neg A$, which does not provide evidence for $A \wedge \neg A$. The preface paradox is one such example.[67]

7 Coniglio and Figallo-Orellano: Categorically Non-Deterministic

In their paper, Marcelo Coniglio and Aldo Figallo-Orellano show us many of the details of the model-theory of the paraconsistent logic *mbC*, a paraconsistent logic with a negation operator and classicality operator both with non-deterministic semantics. There are, naturally, many such systems, including the original da Costa C systems. Coniglio and Figallo-Orellano hereafter C&FO) cite plurivalient logics as another example of logics with non-deterministic semantics. I don't think that this is the right way to look at them. Plurivalent logics are logics in which sentences can take a plurality of semantic values, and if one constructs one on top of a many-valued logic, then the connectives are deterministic, in the sense that the collection of values of the inputs determines the collection of the collection of the values of the output uniquely.[68] However, this is tangential to the main concern of

[67] See IC, 7.4.

[68] See Priest (2014c).

C&FO's paper.

This is not the place to comment on the details of their paper, so I will just say this. In the last 60 years or so, most mathematical work in model theory has been on classical logic. C&FO's paper shows that model-theoretic investigations of non-classical logics, such as paraconsistent logics, can be every bit as sophisticated as that of the model theory of classical logic. There is clearly much interesting mathematics to be undertaken here.

Let me, however, make a few philosophical comments on something central to their paper: the use of a classicality operator. Dialetheism is the view that some statements are dialetheias. As such, it is not committed to any particular view about which statements are dialetheic: that will be the concern of particular applications of the view. The major application investigated over recent years has been to the paradoxes of self-reference—so much so, that many people, I believe, think of it simply as a view concerning these paradoxes. It is not: there are many possible applications of the view; the one concerning the paradoxes of self-reference is a very important one, but it is by no means the only one. It has never even seemed to me to be the most ungainsayable. That honour surely belongs to the application concerning legal contexts, where the ability of legislatures to make things true by fiat is transparent.[69] In the end, in may not be the most profound application either. Applications concerning the limits of thought/language[70] are perhaps more so.

So suppose that one does not endorse a dialetheic solution to the paradoxes of self-reference. Then there is no reason why an appropriate paraconsistent logic should not have a classicality operator, $\circ$. The major objection to this is that, assuming that negation, $\neg$, behaves is a reasonable fashion, one may define the operator $\dagger A$ as $\neg A \wedge A\circ$, and $\dagger A$ will behave as does Boolean negation.[71] This invites the thought that $\neg A$ is not *really* negation, but some other strange operator; consequently, things of the form $A \wedge \neg A$ are not really contradictions, but something else. If the negation symbol has non-deterministic semantics, there is weight to this thought. For the semantics of natural language is compositional: semantic values of wholes are determined by semantic values of parts. This is how we are able to understand complex sentences that we have never heard before. A non-deterministic

[69]See IC, ch. 13.

[70]As found in BLoT and elsewhere.

[71]For some investigation of the matter, see Omori (201+).

negation symbol is a perfectly fine technical device, but just because of its non-compositional nature, it cannot be an adequate account of negation as it is used in a natural language.[72]

If the underlying logic is LP, a logic in which negation has a deterministic semantics, however, this thought is unfounded. In this logic, the truth table for negation may be written as:

A	$\neg A$
t	f
b	b
f	t

Where t is *true only*, f is *false only*, and b is *true and false*. The truth table for † is then:

A	$\dagger A$
t	f
b	f
f	t

Thus, the very semantics of † is *predicated* on the assumption that some things may be true and false, and so of dialetheism. Moreover, ¬ toggles between truth and falsity, just as one should require of negation; † does not. So it is † that is not really negation, and $A \wedge \dagger A$ that is not really a contradiction. The fact that many logicians have taken classical logic to provide the correct theory of negation does not, of course, make it so—any more than the fact that many physicists took classical mechanics to be the correct theory of motion made it so. Both theories could be thought of as true only in virtue of an inadequate diet of examples, as Wittgenstein put it.[73]

If one does endorse a dialetheic solution to the paradoxes of self-reference, matters are less straightforward. For one can then formulate a liar sentence, L, of the form $\dagger T \langle L \rangle$; and if one takes the T-Schema to be formulated with a detachable conditional, this generates triviality.[74] One then has to hold that Boolean negation, and so †, is semantically defective in an appropriate

[72] See Priest and Routely (1989), §2.2.

[73] E.g., *Philosophical Investigations*, §593.

[74] I note, however, that if it is formulated with a non-detachable conditional, such as the material conditional of LP, this is not the case, See Priest (2017), §8.

sense. In that way, the view agrees with that of intuitionists, who also take Boolean negation to be semantically defective—though for quite different reasons. Nor does the classical logician have any reason to feel smug about this. For, together, a truth predicate which satisfies the T-Schema and a negation which satisfies the conditions of Boolean negation, deliver triviality. One cannot have both. And there is no doubt which of the T-Schema and Explosion is the more counter-intuitive. Moreover, to insist that satisfying the rules of Boolean negation is sufficient to *guarantee* a logical operator a meaning manifests a naivety about meaning which *tonk* should have disposed of once and for all years ago.[75]

8 Cotnoir: How to Have your Doughnut and Eat it

Aaron Cotnoir's paper poses a problem concerning points of topological discontinuity. I entirely agree that the problem is a puzzling one, and for the reasons he gives.

He considers a consistent solution to the problem, according to which multiple points can be co-located, and a solution according to which the point of tear is a gluon of the whole which is being torn. He argues that the consistent story is preferable: the gluon story has too many moving parts. I am inclined to agree with him on this. Gluon theory was not designed, after all, to solve this problem. It seems to me, however, that there is a solution which is preferable to both. To see what this is, let us consider the most simple case.

Take a one-dimensional continuum, C, and mark a point, p, on it. Now, tear the continuum at p into a left part, L, and a right part R. Let l be the right hand end of L, and r be the left hand end of R. After the tear, where is p? Then there four possibilities:

- p is l and not r
- p is r and not l
- p is l and r
- p is neither l nor r

[75] For a fuller discussion, see DTBL, ch. 5.

As Cotnoir points out, the first two cases are implausible. Intuitively, the situation is symmetric, and these answers are not. So we are left with the third and fourth cases.

Solutions along both lines are possible. In Case 3, $l = p = r$. Since, patently, l is not r, we require a non-transitive account of identity. Such is entirely possible, and it does not have to be connected with gluon theory.[76] However, such a solution seems to have an insuperable difficulty, as Cotnoir, in effect, points out. After the tear, L and R are disjoint. But this cannot be if p belongs to both.[77] So we are left with Case 4. It would seem bizarre to suppose that p has gone somewhere else. So it must have ceased to exist.[78] One might, I suppose, ask where it has gone. But that would be a rather silly question. It hasn't gone anywhere. The tearing just destroyed it.

Cotnoir's solution is rather different.[79] He suggests that p was, in fact, the two points, l and r, all along. They were just co-located before the tearing. The fact that two point-sized things can be located at the same point obviously poses problems, akin to those of the medieval problem of how many angels there can be on the head of a pin. If they are in the same place, how can they be distinguished? Indeed, the solution seems to reproduce the element of arbitrariness. How come l went to the left, and r went to the right, instead of vice versa? Indeed, how come they didn't both go the same way? And why suppose there were two if one will do the job?

One might, I guess, ask where l and r came from if they weren't there originally. But that question seems to make no more sense than the question of where p went to. They didn't come *from* anywhere: the tearing simply created them. And in any case, one may ask of Cotnoir's solution the reverse question: how come l and r were there in the first place?[80]

[76]If p is an inconsistent object (see below), we can apply the appropriate paraconsistent machinery. See Priest (2010a).

[77]In this way the fission involved here seems to be unlike the fission of an amoeba, where there appear to be no similar considerations.

[78]This does not mean that it has become a non-existent object, though that is a theoretical option. While we are on the issue of non-existent objects, a minor comment about what Cotnoir says about them in the context of gluon theory. He says (§1.2) that if $u \neq u$ then u does not exist, since $\neg\exists x\, x = u$. No. What follows is that $\neg\mathfrak{S} x\, x = u$. The particular quantifier is not existentially loaded. That u does not exist would be $\neg Eu$, where E is the one-place existence predicate. See ONE, §P7.

[79]Actually, he doesn't address this case. I extrapolate from what he says about the others.

[80]Cotnoir's answer (§2.2), if I understand it, is that the number of co-located points is

Let us call the above scenario the *fission case*. We can run this backwards. We start with L and R, and then join l and r at p, which unites L and R into C. Let us call this the *fusion case*. Since this is simply the fission case running backwards, the situation is the same in reverse. In particular, at the joining, l and r go out of existence, and the distinct point p comes into existence.

Let us now turn to the first major case that Cotnoir considers. This is where a sphere is progressively deformed until the point at the top, a, reaches the point at the bottom, b, that being the only place where the two halves of the deformed sphere are joined. The two halves are then torn apart at this location.

The phase of the progression up to the meeting of a and b is essentially the fusion case. The material to which a and b are attached is different from the linear continuum case, since it is three dimensional. However, we still have two distinct points, a and b, which move together and coalesce. Hence, a and b then go out of existence and p comes into existence, being the single point that joins the two halves of the deformed sphere. The next phase of the progression is essentially the same as the fission case. A tear is made at p generating two distinct objects. Thus, p goes out of existence, and two new points come into existence, one on each half, marking the sites of the tear. Note that there is no reason why these two points, a' and b', should be the same as a and b. They can be entirely new. We do not have to worry, as in Cotnoir's case, as to why a went one way, and b went the other. a' *just is* the point on, say, the left hand half, and b' just is the point on the other half.

Let us now turn to Cotnoir's second major example, the sphere deforming into a torus. The first phase of the progression is the same as in the previous case. a and b move together, and go out of existence when the point of tear-to-be, p, comes into existence. The difference comes at the next stage. p goes out of existence, but instead of two points coming into existence, a continuum of points come into existence—those forming the inner circumference of the torus. The fact that more points come into existence than before changes nothing. It still cannot be the case that p is identical to each of these, or the torus would be joined at the middle. And we do not have to suppose, as for Cotnoir, that the point of singularity actually housed a continuum

determined by the potential boundaries the location could be involved in. If this is so, it might seem more plausible to say that the points are there potentially, and are actualised (come into existence) only when the boundary is formed.

of points before the tear (exacerbating the angel problem). Of course, one can ask why so many points came into existence in this case. The answer is simple—and is essentially the same as in the two-point case. A continuum of points is required to deliver the integrity of the post-tear structure. Since, *ex hypothesi*, that is the structure that comes into being, those are the points that come into being. Or, to put it another way: if some of the points did not come into being, we would not have a torus, but a torus missing some bits. If it is a torus that comes into being, so, therefore, must those points.

I have now explained what I take to be a better solution to Cotnoir's puzzle. It seems to me to have none of the problematic features of either of the solutions he considers.

The observant will have noticed that paraconsistency and dialetheism have nothing to do with this solution. Does this mean that they are absent from the scenario? Not necessarily. Come back to a pre-tear structure, say the linear continuum, C. L and R were still there, just united. One may therefore ask whether p belonged to L or R before the split. p is a boundary between the two, and boundaries are—almost by definition—contradictory objects, both separating and joining the things of which they are the boundary. It therefore seems natural to suppose that p is symmetrically poised, in both L and R, but in neither. It joins them because it is in both; and it separates them because it is in neither. So, if $x < p$ then x is consistently in L; if $x > p$ then x is consistently in R; and p is inconsistently in both.[81]

9 Dicher and Paoli: ST, LP, and All That

Bogdan Dicher and Francesco Paoli take us into the world of Cut-free logics. Before I turn to their paper, some preliminary comments. In standard many-valued logics, there is a single set of designated values. The technique of having a different set of values for the premises and conclusion(s) is an interesting one, and allows for a logic which can invalidate principles such as Identity and Cut, which standard many-valued logics validate.

The use of such semantics by Cobreros, Egré, Ripley, and van Rooij to deliver a logic ST, which may provide a solution to paradoxes such at the liar, is an intriguing one. Proof-theoretically, one may obtain a consequence relation which delivers the same valid inferences as classical logic. However,

[81]For a formal model, see IC, 11.3. For Cotnoir's own take on inconsistent boundaries, see Cotnoir and Weber (2015).

when there are axioms or rules of inference for distinguished predicates, such as the T-schema, the failure of Cut blocks the argument to triviality, since one may have things of the form $\Rightarrow A$, $\Rightarrow \neg A$, $A, \neg A \Rightarrow B$, but not $\Rightarrow B$.

Now, first, it has always seemed to me that this is not a way of avoiding dialetheism with respect to the liar paradox. Given self-reference and the rules endorsed for truth, one can construct a liar sentence, λ, in the usual way, and establish that $\Rightarrow \lambda$ and $\Rightarrow \neg\lambda$ (and so $\Rightarrow \lambda \wedge \neg\lambda$). In other words, the principles for truth establish contradictions; so we have dialetheism. From this perspective, the logical machinery is just a different way of endorsing dialetheism and blocking to Explosion.

Next, since in this logic $A, \neg A \Rightarrow B$, it is not paraconsistent, at least according to the standard definition. But since the argument from a contradiction to triviality is blocked, the conclusion one might well draw from this is that the standard definition is wrong (or at least, works only when Cut is present). A more adequate definition might be that a consequence relation, $\Rightarrow$, is paraconsistent if there are theories, $\mathcal{T}$, and sentences, A, B such that $\mathcal{T} \Rightarrow A$, $\mathcal{T} \Rightarrow \neg A$, but not $\mathcal{T} \Rightarrow B$.[82]

Third, is this logical machinery the best one for a dialetheist to endorse? Perhaps. The machinery obviously has its attractions. As far as the liar paradox goes,[83] it is very neat. Whether it works so well for other arguably dialetheic areas, such as law, motion, the limits of thought/language is an issue that still needs to be addressed. The main problem is, of course, that the logic is very weak. All the classically valid inferences might be available, but without Cut, much apparently unproblematic reasoning (for theories with non-logical axioms/rules) is unavailable. The same is true of the logic LP, of course. (I pointed this out in the very first paper I wrote on the subject.[84]) Over the years, many ways of overcoming this weakness have been investigated, such as augmenting the logic with a detachable conditional, and employing default reasoning.[85] Whether the Cut-free approach can do as well or better in the matter is an issue so far unaddressed, and far too large an issue to take on here.

[82]The alternative definition of 'paraconsistent', together with the relevance of Cut (transitivity) was already piointed out in §1 of Priest and Routley (1989a).

[83]And maybe the set-theoretic paradoxes as well, though this puts the question of "classical recapture" at centre stage.

[84]Together with ways of addressing the matter. See Priest (1979), §4.

[85]More recently, I have explored the matter of classical recapture in mathematics in terms of mathematical pluralism. See, e.g., Priest (201+b).

This brings us to Dicher and Paoli (hereafter, D&P).[86] In axiomatic presentations of logic, there is a standard distinction between rules that are truth-preserving (like *modus ponens*) and rules that are merely validity-preserving (such as Necessitation). The latter are fine if one is just trying to generate the logical truths, but if one is using logic as an organon of proof in general, the former are required. D&P point out, correctly, that exactly the same distinction, with exactly the same point, can be made with respect to the rules of a sequent calculus. In the light of this, they show that the sequent calculus given by the authors of ST is incomplete with respect to the semantics, and provide an extended sequent calculus, LK^{-}_{INV}, which is complete. This is an insightful piece of work.

Turning to its philosophical consequences, D&P note that, modulo a very natural notion of when distinct formulations deliver the same logic, LK^{-}_{INV} and LP are the same logic.[87] Given this, any rivalry between ST and LP would seem to disappear. As they point out, one might contest the notion in question. Whether it is appropriate to do so, and, if so, what the upshot is, I will leave for the authors of ST to determine.

However, D&P have another card in their hand. Even supposing that ST and LP are distinct, it remains the case that the language of ST is incomplete. Just as one might expect there to be logical constants to represent the values 1 and 0, namely $\top$ and $\perp$, one should expect there to be one that represents the value 0.5. A constant behaving as does λ fits the bill nicely. In other words, one does not need the liar paradox to deliver dialetheism: it is built into logic itself.

I take all this to be an elegant way of making much more precise my initial reaction to ST—that it is very much dialetheism-friendly.

10 Dunn and Kiefer: Heeding Firefighters

Michael Dunn is a veteran warrior of relevant logic. Our interest in relevant logic brought us together in the early 1980s, and we have been friends since

[86] One small side-comment. D&P mention the connection between Gödel's incompleteness theorems and semantic closure (and specifically the deployment of a naive truth predicate). This is certainly a connection I have stressed before (e.g., IC, ch. 3). However, the points made concerning the import of the theorems for dialetheism can be made without appeal to a truth predicate. See IC2, ch. 17.

[87] The result was proved in Pynko (2010). The connection between LP and LK minus Cut was, in fact, already noted in *IC*, p. 78.

then. In their paper, he and Nicholas Kiefer present a problem about conflicting information in the context of a paraconsistent logic. Before we turn to this, a couple of preliminary comments.

Dialetheism is one application of paraconsistent logic; but of course it is not the only one, as I have often pointed out. Handling information for which one cannot guarantee consistency is another.[88] Such an application might well suggest that, in the semantics of a paraconsistent logic, the values *true*/*false* should be thought of as *informed as (told) true*, and *informed as (told) false*. Such an interpretation has appealed to many of those, such as Nuel Belnap and Dunn himself, hailing from the US Pittsburg relevant logic group (as opposed to the Australian Canberra group).

This interpretation has well known problems, though. The first concerns conjunction. One is often given the information that A and the information that $\neg A$, where one would would not want to claim to have been given the conjoined information, as is required by the semantics of relevant logic. Indeed, the situation given by Dunn and Kiefer (hereafter, D&K) is exactly an example of such a thing.[89] The poor person fleeing from the burning hotel room is not at all inclined to infer that there is one exit which is both to the left *and* to the right. The informational understanding of semantic values motivates much more naturally a discussive paraconsistent logic, where conjunction-introduction fails.[90] A second problem besets disjunction: one is often informed that $A \vee B$ without being informed either that A or that B. But on the favoured Belnap/Dunn interpretation, if $A \vee B$ is "told true" so is either A or B.

Moreover, there is nothing about the more standard reading of the semantic values as *true* and *false* which commits one to dialetheism. These are values that a sentence has in an interpretation (or even a world of an interpretation). There is nothing to imply that interpretations in which things are both true and false represent *actual* situations. In logic, one reasons about many sorts of situation: actual, possible, and maybe impossible too. A non-dialetheic advocate of relevant logic may hold that the interpretations in which things are both true and false are of the latter kind.[91]

With these preliminary remarks out of the way, let us turn to D&K's Paradox of the Two Firefighters. Actually, I wouldn't call this a paradox.

[88] See, e.g., Priest and Routley (1989b), §4ζ, Priest (2002), §2.2.

[89] As is the paradox of the preface. See IC, §7.4.

[90] See Priest (2002), §4.2.

[91] See Priest (1998), p. 414.

There is nothing paradoxical about the situation in which the escaper-to-be finds themself. It could be all too real. Still, there is certainly a puzzle here. What should the person do in the face of the contradictory information, and why? Intuitively, without information, the person should chose from the three directions at random. If only one firefighter speaks, they should go that way. If two speak, they should chose between the two directions at random. The question is, then, why? D&K suggest two possible solutions to the problem, one of which uses a paraconsistent logic, and one of which uses classical probability theory.

Although D&K do not set things up this way, this is really a problem of decision theory ('there is nothing like the pressure of needing to decide which hallway to take to avoid being burned to death', §3). So let me give what seems to me to be the most natural decision-theoretic solution to the puzzle. (To what extent this is preferable to D&K's two solutions, I leave the reader to consider.)

Let G_L, G_R, and G_S be, respectively: go left, right, and straight on. And let F_L, F_R, and F_S be, respectively: find an exit left, right, and straight on. Let $pr(A)$ be the probability of A; let $val(A)$ be the value of A; and let $\mathcal{E}(A)$ be the expectation of A. Then:

- $\mathcal{E}(G_L) = pr(F_L).val(F_L) + pr(\neg F_L).val(\neg F_L)$
- $\mathcal{E}(G_R) = pr(F_R).val(F_R) + pr(\neg F_R).val(\neg F_R)$
- $\mathcal{E}(G_S) = pr(F_S).val(F_S) + pr(\neg F_S).val(\neg F_S)$

Let the value of finding a door be v, and the value of not finding one be $-v$, where v is some positive real.[92] And prior to any information, we can assume that the probabilities of each of the Fs is the same, namely, $p > 0$ (or else prepare to die!). Thus:

- $\mathcal{E}(G_L) = \mathcal{E}(G_R)$=$\mathcal{E}(G_S) = pv + (1-p). - v = v(2p-1)$

Since these actions all have the same expectation, one should chose between them at random.

Now, suppose that we have only the information given by the *go-left* firefighter, then:

[92]Perhaps one should not assume one is the negative of the other; but nothing much hangs on this assumption, and it keeps matters simple.

- $\mathcal{E}(G_L)$= $1v + 0. - v = v$
- $\mathcal{E}(G_R)$=$\mathcal{E}(G_S) = 0v + 1. - v = -v$

Clearly, going left has the highest expectation, and so one should do that.[93] Of course, given only the the information given by the *go-right* firefighter, the situation is symmetric, so:

- $\mathcal{E}(G_R)$= $1v + 0. - v = v$
- $\mathcal{E}(G_L)$=$\mathcal{E}(G_S) = 0v + 1. - v = -v$

So we come to the case with inconsistent information. One of the possibilities that D&K consider is that of aggregating probabilities; but we can aggregate expectations directly instead. We may suppose that each of the two bit of information is just as good as the other, and so average them out. This gives:

- $\mathcal{E}(G_R)$= $\frac{1}{2}(v - v) = 0$
- $\mathcal{E}(G_L)$= $\frac{1}{2}(v - v) = 0$
- $\mathcal{E}(G_S) = \frac{1}{2}(-v - v) = -v$

Clearly, going left and right have equal expectations, and a better expectation than going straight on. So one should chose between these two at random. Hence, all our pre-theoretic intuitive judgments have been vindicated.[94]

Finally, I note that this solution has nothing to do with paraconsistent logic.

[93]Note that provided that $p < 1$, $v > v(2p - 1)$, so the information has increased the expectation of going left. An increase of information does not have to guarantee that an expectation goes up. Thus, suppose that the firefighter says that there is one exit, and it is to the left, but that it is probably closed. Then the probability of F_L becomes 1, but the value of finding it, and so the expectation of going left, drops. One would, none the less, prefer to have this information than none.

[94]This analysis treats the information given as dependable. We may take into account the possibility that it is not so with a simple dominance argument. If it is dependable then, as shown, one should go left. If it is not, each direction is equally good. So going left is the dominant strategy.

11 Égré: Respectfully Yours

There is a time-honoured strategy for resolving an apparent contradiction: draw a distinction. And this is clearly exactly the right thing to do in many cases. I say truly, 'it is 5pm and it is noon'. That sounds like a contradiction, but in fact, what I mean is that it is 5pm in London, and noon in New York. In this case, the family of parameters required to disambiguate is discrete. (There is a finite number of time zones.) But the family can equally be continuous. Thus, one might say that something is red to degree 0.6, and not red to degree 0.4, the parameters being the real numbers. In such a case, one may fairly speak of the property as coming by degrees.

This raises the question of whether all apparent dialetheias may be resolved by the strategy of paramaterisation. According to Paul Égré, they can:[95]

> I argue that this relation [equivocation more subtle than lexical ambiguity]... underlies the logical form of true contradictions. The generalization appears to be that all true contradictions really mean "x is P is some respects/to some extent, but not in all respects/not to all extent".

This raises two question: [α] Are all apparent dialetheias produced by predicates subject to parameterisation? [β] In cases where they are, does this resolve the contradiction involved? Let me address these questions in turn.

11.1 The Variety of Dialetheias

Égré offers no argument for a positive answer to [α], and one is necessary. For this certainly does not appear to be the case. The putative examples of dialetheias that have been offered include:

1. Paradoxes of semantic self-reference, such as the liar.[96]

[95] Abstract. In one place, he appears more circumspect, saying that he is concerned only with contradictions with respect to vague predicates. ('In this paper I propose to [address]... the semantic treatment of contradictory sentences involving vague predicates' (Introduction).) But the rest of the paper eschews this qualification. So I presume that he thinks that all *prima facie* dialetheias involve vague predicates. At the very least, nothing in the paper shows any awareness of other possibilities.

[96] E.g., IC, ch. 1.

2. Set theoretic paradoxes, such as Russell's.[97]

3. Paradoxes concerning Gödel's theorems (involving the sentence 'this sentence is not provable').[98]

4. Contradictions concerning the instant of change, and, more generally, motion.[99]

5. Legal contradictions concerning inconsistent legislation.[100]

6. Paradoxes concerning the limits of language (or thought), in which one appears to have to describe the ineffable.[101]

7. Statements concerning the unity of objects.[102]

8. Statements connected with multi-criterial terms.[103]

9. Statements concerning the borderline area of vague predicates.[104]

8 and 9 relate to question [β]; and Égré argues that 1 is a special case of 9. I will return to these matters in a moment. This is hardly the place to comment at length on the others, but let us review them briefly.

2 seems to have little to do with matters of respect or degree. The membership of a set does not come by degrees: it is an all-or-nothing matter. Note that the sets involved in the paradoxes in question are not fuzzy sets, as the set of all males might be. They are perfectly crisp mathematical objects deploying only pure-mathematical predicates. Similar points apply to 3. This concerns what is mathematically provable. Mathematical proof is deductive. The strength of a deductive argument does not come by degrees, as does that of an inductive argument.

For 4, just consider the situation at an instant of change, e.g., when I am leaving a room, and symmetrically poised between being in and not in it. Actually, this particular case could well be thought of as involving vagueness,

[97]E.g., IC, ch. 2.
[98]E.g., IC, ch. 3.
[99]E.g., IC. chs. 11, 12.
[100]E.g., IC, ch. 13.
[101]E.g., BLoT.
[102]E.g., ONE, Part 1.
[103]E.g., IC, §4.8.
[104]E.g., Priest (2010b).

and so degrees, since both I and the room are extended (and vague) objects. But this is not essential to the example, which applies just as much to point-particles crossing a two-dimensional boundary. (See IC, p. 161.) As another example, consider the point of midnight between Monday 1st and Tuesday 2nd. There is no vagueness in either Monday, Tuesday, or the instant of midnight. All are, or are bounded by, precise points.

Examples of 5 occur when there is a law of the form: people in category X may do Z; people in category Y may not do Z. Someone, a, then turns up who is in both categories X and Y. Now, the categories may themselves be vague, but the membership of a in X and Y can be as determinately true as one might wish. Note, also, that in legal matters, there is room for a qualification of respects. Thus, something may hold in one jurisdiction but not in another. But in the case at hand, both clauses may be taken as parts of a single law in a single jurisdiction.

Examples of kind 6 depend on arguments to the effect that something or other is ineffable. There are many such arguments, and different philosophers have endorsed different ones. In some of these, one might try to invoke a distinction of respect, though rarely does this seem to succeed. Consider just one example: orthodox Christianity says that God is so different from his creatures that it is almost impiety to suggest that human categories apply to him. Yet theologians say much about God—with human categories. (What else do they have?) In response to this, theologians of a *via negativa* persuasion have suggested that one can say nothing *positive* of God. One can say only that he is *not* this and *not* that. But theologians seem do say an awful lot of positive things about God (such as that he is omnipotent, omniscient, etc). Indeed, even the claim that (only) negative things can be predicated of God is positive. But in any case, if we take to heart the claim that God is literally ineffable, then one can say *nothing* about God—not even that. There is no room for respects or degrees. Even to say *a little bit* is to say something.

Finally, 7. This arises when we are forced to recognise that something (a gluon) both is and is not an object. As such, this can be seen as a special case of 6. For if something is not an object, one can say nothing of it, so is it is ineffable: to say 'a is such and such' is to treat a as an object. But for the same reason, when one says something, this requires the object of predication to be just that—an object.

11.2 Dialetheism and Differences of Degree/Extent

So let us turn to question [β]. But first, a preliminary comment about this kind of case. To make their point, Égré, and many of those like him who approach philosophical issues via empirical linguistics, appeal to what (many/most/some) people are wont to say. Such a form of argument must be treated with great care (as Égré, in effect, notes), since people will say all kinds of false things: 'the sun rises in the morning', 'the sky is blue', 'white people have often oppressed black people' (the white people are sort of pink, and the black people are various shades of brown). What people say might provide some *prima facie* evidence for the truth of what is said. But in the end, this cannot replace looking at the evidence/arguments for its truth.

That said, let us start with examples of case 9. These arise when one has a vague predicate, say *red*, and something is on the borderline between red and not-red. Égré suggests that a vague predicate, P, depends on a property that comes by degrees. Let us write the degree of a's being P as $|a|_P$, where this is a member of some closed interval of non-negative reals, X. Égré holds that there are $\theta_0, \theta_1 \in X$ (determined by context), with $\theta_0 < \theta_1$, such that a is completely P if $\theta_1 \leq |a|_P$; a is partially P (or P in some, but not all, respects) if $\theta_0 < |a|_P < \theta_1$; and a is not at all P if $|a|_P \leq \theta_0$. Borderline cases are those as in the middle class.

There are already issues here to do with higher order vagueness, but let us set such worries aside.[105] The picture so far does not avoid dialetheism, for the simple reason that we have so far said nothing about $\neg P$, which is equally a vague predicate. Given $|a|_P$, where is $|a|_{\neg P}$? Could both be $\geq \theta_1$? We need a theory of negation.

Égré does not address the matter in the paper, but he refers us to his work with Cobreros, Ripley, and van Rooij, where they endorse the logic ST. Semantically, this is a three-valued logic, with values, 1, 1/2, and 0. Negation maps 1 to 0, vice versa, and 1/2 to 1/2. We may take Pa to have the value 1 if $\theta_1 \leq |a|_P$; the value 1/2 if $\theta_0 < |a|_P < \theta_1$; and the value 0 if $|a|_P \leq \theta_0$. So Pa is partially true iff both Pa and $\neg Pa$ are "half true".

Now, there are a number of things to be said about this. First, there are three-valued logics for negation in which both Pa and $\neg Pa$ have the value 1. For example, there are fuzzy relevant logics where both can have the value 1

[105] All accounts of vagueness have issues with higher order vagueness. See the introduction the Keefe and Smith (1999). I have had my say on matter in Priest (2010b).

(at a world).[106] Theories of negation are many. Hence, we need an argument for Égré's theory.

Next, note that there will, in general, be differences of degree *within* the three categories. Pa and Qa may take the value 1, even though $|a|_P > |a|_Q$. It would seem that for every a such that $|a|_P$ is less than the maximum value (if, indeed, there is one) Pa is true in some respects, but not all. To call only those in the middle category *partially* true, or true to some extent, therefore seems to misrepresent the situation:

It is more plausible to interpret the trichotomy as *just true*, *true and false*, and *just false* (as in LP).[107] The question is whether the value of $|a|_P$ is such as to make both Pa and $\neg Pa$ assertible—which would seem to be the case if one can take the empirical data as indicative of this.[108] We are back with unvarnished dialetheism.

Let us now turn to examples of case 8. These arise where a predicate can have different criteria of application, which can fall apart. Thus, consider the predicate 'has a temperature of $x°$ absolute'. There are different ways of verifying a statement of this form. We may measure temperature with a mercury thermometer, an electro-chemical thermometer, the frequency of black-body radiation emitted, and so on. The method used may depend on both the object in question and on how hot it is. But in some cases, different measuring devices may be used to measure a temperature of the same thing. (Thus, the temperature of sea water can be measured by both a—correctly functioning—mercury thermometer and a—correctly functioning—electro-chemical thermometer. And normally in such cases, these will give the same result (to within experimental error). But such determinations way well come apart. In such a case an object can have a temperature of $x°$ and not of $x°$. Indeed, such cases are not unknown in the history of science. A plausible historical concerns the notion of the angular size of bodies in seventeenth-century optics.[109]

Now, it might well be thought that predicates of the form 'has a temperature of $x°$' are examples of what Égré calls polar opposites, and so may be

[106] See INCL, ch. 11.

[107] Alternatively, if one is really serious about degrees, it makes more sense to move to a fuzzy logic—which raises quite different issues.

[108] Cobreros, *et al* (2012), (2013), call things in category 1 strongly assertible, and things in category 1/2 weakly assertible. These are terms of art. The question is simply whether $|a|_P$ is such as to make both Pa and $\neg Pa$ true enough to be assertible (in the context).

[109] See Maund (1981), pp. 317f.

handed by some form of aggregation (Égré's f), described in §4. But this would be to misunderstand how such predicates work. If we found ourselves in this situation of conflict, and assuming that this was not due to experimental error, but a quite systematic phenomenon, we would most certainly not compute the degree to which an object has a temperature of x° by taking a weighted average. What would happen is that the notion of temperature would bifurcate into two, one determined by each of the criteria.[110]

Nor should the fact that such conceptual fission takes place show that there was no dialetheia. The new concepts are (one would hope) consistent. But it remains the case that statements involving the old concept were inconsistent. The contradictions are verifiable, and so true—even if the verifications are never actually performed.[111] And of course, we may well have concepts of this kind for which we never discover the inconsistency, and so which never undergo conceptual fission.

Finally, let us come to examples of kind 1, the liar paradox, concerning a sentence, L, or the form $\neg T\langle L\rangle$. Égré argues that the truth predicate is a vague predicate, and so the contradiction of the liar is defused by the general considerations concerning vagueness. There are a couple of things to be said about this matter.

The first is that it is not clear that the truth predicate is a vague predicate. Why not take T to be a perfectly crisp predicate? Even if P is itself a vague predicate, we may hold that $T\langle Pa\rangle$ takes the value 1 or 0 depending on whether the degree to which a has the property P is greater or less than some (contextually determined) cut-off point. In other words, T is a crisp but indexical predicate.

But suppose that the value of Pa is identical with that of $T\langle Pa\rangle$. Then if A has a classical value, so does $T\langle A\rangle$; and if A has a non-classical value, so does $T\langle A\rangle$. But what of $T\langle L\rangle$? To argue that this has a non-classical value because L does is clearly question-begging. *Any* putative solution to the liar paradox may say that things must be *thus and such*, or a contradiction will

[110]Conceptual fission of this kind is well known in the history of science. Thus, the notion of *mass* in Newtonian dynamics bifurcated into *rest mass* and *inertial mass* in the dynamics of Special Relativity. See Field (1973).

[111]One might take such predicates to have multiple referents, which can be used to parameterise the predicate. However, be the referents many, the sense is one; and each verification is sufficient to apply the predicate. (See Priest and Routley (1989c), §2IIi.) Moreover speakers who used the term *mass* most certainly did not *mean* 'rest mass or inertial mass'.

arise. What is necessary is an independent argument that things are *thus and such*, or the move is entirely *ad hoc*.

Next, and most importantly, the semantic value of the liar sentence is, in a sense, neither here nor there. The liar paradox is generated by an argument. Here is one way of putting it, in sequent-calculus form:

$$\begin{array}{rcll}
T\langle L\rangle & \vdash & T\langle L\rangle & 1\\
T\langle L\rangle, \neg T\langle L\rangle & \vdash & & 2\\
\neg T\langle L\rangle & \vdash & \neg T\langle L\rangle & 3\\
L & \vdash & L & 4\\
T\langle L\rangle & \vdash & L & 5\\
 & \vdash & \neg T\langle L\rangle, L & 6\\
 & \vdash & L, L & 7\\
 & \vdash & L &
\end{array}$$

1 is the identity sequent. 2 and 3 are applications of the rule for negation. 4 holds since $\neg T\langle L\rangle$ just is L. 5 is an application of the T-schema. 6 is the rule for negation again. 7 is just the identity of L and $\neg T\langle L\rangle$ again; and 8 is contraction. Similarly:

$$\begin{array}{rcll}
T\langle L\rangle & \vdash & T\langle L\rangle & 1\\
T\langle L\rangle, \neg T\langle L\rangle & \vdash & & 2\\
\neg T\langle L\rangle & \vdash & \neg T\langle L\rangle & 3\\
L & \vdash & L & 4\\
L & \vdash & T\langle L\rangle & 5\\
L, \neg T\langle L\rangle & \vdash & & 6\\
\neg T\langle L\rangle, \neg T\langle L\rangle & \vdash & & 7\\
\neg T\langle L\rangle & \vdash & & 8\\
L & \vdash & & 9\\
 & \vdash & \neg L & 10
\end{array}$$

But Égré endorses the T-schema (§5.1), and all the other steps are valid in the logic ST, which is his preferred logic, given the analysis in his paper. In the logic one can even conjoin the conclusions to infer $\vdash L \wedge \neg L$.[112] So the liar contradiction follows from premises and rules that Égré endorses. One should therefore accept it. Dialetheism.

His analysis will not, then, defuse the contradiction of the liar paradox.

[112] It is also the case that $A, \neg A \vdash B$, for arbitrary A and B; but one cannot infer $\vdash B$, since Cut in not valid.

12 Ferguson: Collapsing in Unusual Places

I discovered the Collapsing Lemma in the late 1980s, when working on a technical problem concerning minimally inconsistent models.[113] Later, I discovered that it had quite different applications in the construction of inconsistent models of arithmetic and set theory. The models were both technically interesting in their own right, and had important philosophical ramifications—for example, connected with Gödel's Theorems, and various aspects of set theory.[114] In his paper, Thomas Ferguson notes these things, and goes on to establish a number of interesting results concerning the possibility or otherwise of extending the Collapsing Lemma to machinery beyond that of LP.

Speaking generally, the Collapsing Lemma depends on the truth functions and quantifiers involved behaving monotonically, in a certain sense. Namely, in Ferguson's notation, whenever an input changes from $\mathfrak{b}$ to $\mathfrak{t}$ or $\mathfrak{f}$, the output is never changed from $\mathfrak{t}$ to $\mathfrak{f}$, or vice versa. That is, consistentising an input cannot change a classical output. The result can therefore be extended beyond the bounds of first-order logic to other logical machinery which is monotonic—for example, second-order quantifiers and modal operators.[115] However, the Lemma will fail once one is dealing with non-monotonic connectives/quantifiers. In particular, there is no version of the Lemma which will work for a many-valued truth-functional logic with a detachable conditional, as Ferguson nicely shows; and I have no idea of how one might turn the trick for a logic of some other kind.

Nice as it would be if this were possible, I don't see that it is a serious problem if it is not. The Lemma was always part of an investigation to see what could be done with classical machinery in a paraconsistent context. Classical logic elects to work within the framework of conjunction, negation, the particular quantifier, and the things that can be defined in terms of these—and in LP, these are all monotonic in the appropriate sense. In particular, classical mathematics works within this framework. Hence, a limitation of the Collapsing Lemma implies no limitation of the investigation of paraconsistent versions of classical theories. The mantra of paraconsistent

[113] See Priest (1991). I subsequently found out that a similar result had been proved by Mike Dunn (1979) some years earlier.

[114] See IC2, chs. 17, 18, and Priest (2017), §11.

[115] See Priest (2002), §§7.2, 7.3. On many-valued modal logic in general, see INCL2, ch. 11a.

logic with regard to classical logic/mathematics has always been: we can do anything you can do—and a lot more besides![116]

13 Ficara: Priest as a Hegelian (or Hegel as a Priestian)

It has always seemed to me that the most salient and ungainsayable dialetheist in the history of Western philosophy, between Aristotle's wildly influential but fatally flawed attack on the view in the *Metaphysics*[117] and contemporary times, is Hegel. I know that suggesting that Hegel was a dialetheist is wont to provoke fits of apoplexy in a number of Hegel scholars, who can see no further than so called "classical" logic.[118] However, I have defended the view elsewhere,[119] and I shall not revisit the issue here.

Elena Ficara's paper concerns a different (though related) possible similarity between my view and Hegel's. Now, the world *logic* gets used in many ways, and I think it is silly to argue about what the right use is. One just needs to be clear about how someone uses the word. Hegel and I use it in somewhat different ways, as will be clear to anyone who compares Hegel's *Logic* and my *Introduction to Non-Classical Logic*. I use the word in the way that most contemporary logicians use it, as being about validity—that is, what follows from what, and why. For Hegel, logic is what is covered in his *Logic*—roughly, human conceptual thought and its rational evolution. I think that what I mean by logic is pretty close to what Hegel calls *subjective logic* (*Verstandeslogik*). For him, this is the Aristotelian syllogistic of his day.

That point nothwithstanding, and assuming that Ficara's exegesis of Hegel is right,[120] there does seem to be a similarity between, pairs of our terms: (a) *logica utens*, *natürliche Logik*; (b) *logica docens*, *die Logik*; (c) *logica ens*, *das Logische*. The first comprises the norms of a reasoning practice. The second comprises logical theories, as found in logic textbooks. The third is what it is that such theories aim to capture. And as Ficara points out, according to both Hegel and myself, (b) can be revised. In particular, it can

[116]On that matter, see IC, ch.8, esp. 8.5.
[117]See DTBL, ch. 1
[118]Though not Ficara, as I know from many illuminating discussions.
[119]See Priest (1990) and (201+c).
[120]I am no Hegel scholar, and am usually happy to leave the nitty-gritty of Hegel-exegesis to those, like Ficara, who know his work better than I do.

be changed in order the better to bring it into line with (c).[121]

The major difference that Ficara points out concerns this process of revision. For me this is to be done by a very general process of theory-choice.[122] For Hegel, it happens in the process of dialectical thought (undertaken by *Geist*). This is his *Vernunftlogik*.[123]

Concerning this matter, Ficara says (§3):

> While Hegel postulates the idea of a rational logic, i.e. embeds his view of logical revision in a conception of logic as conceptual and philosophical (i.e. self-revising and truth-oriented) analysis of natural language, Priest sticks to the idea of an external operation, which follows the model of rational theory choice among rival logical theories. Logic revision is for Hegel, as we have seen, the result of Hegel's very idea of logic as the analysis of *das Logische*, i.e. of logic as analysis and individuation of forms of truth. Revision, intended as procedure of adjustment between theories and data, is an operation actuated by logic itself. In this sense Hegel's idea of logic does not admit the distinction between pure and applied logic. In non-Hegelian terms, Hegel's logic as rational logic would involve both the construction of a model and reflection on the adequacy of the model.

Now, I think that this matter is, to a large extent, terminological. Hegel uses the word *logic* to encompass rationality in general. I use in for just one aspect of rationality. We both agree that there is a process of rational revision going on; and (assuming that Ficara is right about Hegel) we both

[121] Concerning (a), Ficara says (§2) says that 'one can/also admit that reasoning practices and norms are grounded in metaphysical views about what there is and its nature. Priest implicitly assumes this insofar as he states that inferential norms are based on our view about the meanings of the connectives'. Now there certainly can be metaphysical assumptions inherent in a reasoning practice—think of Dummett on classical logic *vs* intuitionist logic—but I wouldn't say that disputes about the meanings of the connectives necessarily have metaphysical ramifications.

[122] See Priest (2016a). This can be thought of as a sort of reflective equilibrium as Ficara (fn 19) notes. This does not mean that it does not 'admit the metaphysical meaning of logic'. *Logic ens* is the very subject of the deliberative equilibrium. I note also that theories of *validity* are liable to engage with theories of other notions, such as *meaning* and *truth*. In that way, theorising about validity is entangled with a number of the more general issues which play an important role in matters for Hegel.

[123] For a formal account of this, see Priest (201+d).

agree that this involves a dialectic between theory and data.[124]

However, it is certainly true that Hegel and I think of the mechanism of revision quite differently. Hegel's view, I take it, can be accepted only if one endorses Hegel's idealism. This has always struck me as somewhat whimsical view; but I certainly don't intend to defend that claim here. So let me just defend my own view against the problems that Ficara sees.

Ficara notes essentially two problems. The first concerns the fallibility of the data of theory-choice (§3):

> why should the data be a criterion of rationality if they are our intuitions about validity and our intuitions can be wrong?

The second is that the method needs an explanation of the connection between the data of logical theorising and what actually follows from what (§3):

> Priest's theory about logic revision needs to be completed by a conception about the nature of the data, and their relation to truth, or *logica ens.*

Take the second point first. Theorising is an important rational activity. We are aware of some phenomenon, be it motion, time, ethics, language. We wish to understand the whats and whys of it. To do so we construct a theory to account for the phenomenon. Since we are aware of the phenomenon, there are already things we believe about it, or can ascertain directly. These provide the initial data to which the theory answers.

Now, giving reasons is something we all do naturally, though perhaps badly. It is certainly a skill that can be improved if one studies mathematics, law, and doubtless many other subjects. We come to see that sometimes a reason offered really does support a conclusion; and sometimes that it does not. In other words, there seem to be facts about what follows from what. *Logica ens* is the truth of this matter; *logica docens* is a theory about what the truth is. (And the best theory we have at any time is our best guide to what the reality it aims to describe is like.) The initial views we have about

[124] And for what it is worth, it seems to me that the distinction between (in my terms) pure and applied logic makes just as much sense for Hegel as it does for me. A pure logic is simply a bit of pure mathematics, which may be applied for many purposes, or never applied at all—though Hegel is working before the impact of non-Euclidean geometry on mathematics, which finally brought home the distinction between pure and applied mathematics.

what follows from what are data to which the theory must answer. These are the views which characterise—at least initially—the very phenomenon at issue.

Which brings us to the first question. Adequacy to the data is a criterion (though not the only one) of what it is that makes a theory rationally acceptable. (To account for these is, after all, a large part of what the theory was produced *for*.) In that sense, this is a 'criterion of rationality'. However, this does not mean that the data are infallible. In theorisation about any complex issue, data are fallible. That this is so even for empirical data is one of the hard lessons of the 20th Century philosophy of science.[125] Inadequacy to a piece of data is certainly a rational black mark against a theory; but a theory that is strong in other respects can overturn the datum—especially if one can give an independent reason as to why one was mistaken about its truth. This does not undercut the rational use of data: the dialectic between theory and data, in which views about the truth of both can be revised, just is a feature of rationality.[126]

14 Field: Out in Left Field

Which brings us to Hartry Field. Field and I have been discussing logic, and especially solutions to the semantic paradoxes, for many years now; and the discussions have caused me to think much harder about many issues. I have enjoyed the discussions, and learned a great deal from them. I very much value Field's open-mindedness and intellectual honesty. We agree

[125]See, e.g., Chalmers (2014), ch. 2.

[126]Ficara writes (§3) that the criteria for the revision of logical theory:

> are, for Priest, the standard ones of rational theory choice: adequacy to the data, fruitfulness, non-ad hocness etc. For Hegel, the only criterion that orients the critique of logic as theory is *das Logische* as conceptual truth, as the way the connectives *are* and validity *is*. It is a realistic meaning of truth, the correspondence of the logical theories with the logical fact. Interestingly, a correspondence that is already given in the empirical data logic deals with (our intuitions about validity).

Now, I think that there is a certain confusion here. The aim for both Hegel and myself it to get logic as theory right, or at least, better. In that sense, we both want to get theory to correspond to reality. Where we might disagree is on what method will achieve this, and why. In neither case is this correspondence already guaranteed by our intuitions, which, we can both agree, are sometimes wrong, or at least, inadequate.

about much, and especially on the fact that “solutions” to the paradoxes which employ “classical” logic are not viable.

The solutions to the paradoxes that he and I endorse also have much in common. But there is one crucial disagreement: his solution to the paradoxes is to reject Principle of Excluded Middle (PEM); mine is to reject Explosion. His present paper is the latest written contribution to our discussions.[127] As ever, it is rich and insightful—and there is absolutely no way that I do justice to everything it contains here. (To do this would require a piece at least as long!) So I will restrict myself to the most important things. I will take these up in more or less the order in which he raises them, though I collect together a bunch of more minor points in a section at the end.

14.1 K_3 and LP: Duality

Field points out (§1) that there is a substantial duality between K_3 (“middle value” non-designated) and LP (“middle value” designated). He concludes on the basis of this that, setting aside issues of conditionality, the dispute between advocates of the two logics is simply notational. In particular, the LP theorist’s acceptance of A is the K_3 theorist’s rejection of $\neg A$, and vice versa. Perhaps, as far as formal matters go, this is correct. But enforcing this duality appears to have implausible consequences in a wider context, since acceptance and rejection have essential roles to play in other areas, such as action and its rationality. Someone who accepts that it will rain has grounds to take an umbrella. Someone who merely rejects the claim that it will not rain does not have the same ground; for this fact, classical logic having gone, gives them no reason to suppose that it will rain! Similarly, someone who is a dialetheist about the liar paradox, L (such as me), has grounds to write a book advocating L and $\neg L$. Someone who merely rejects both (such as Field) does not. To interpret Kripke’s ‘Outline of a Theory of Truth’ (1975) as advocating both L and $\neg L$, would seem to be an act of gross perversity!

14.2 Conditionals

Let us now turn to the subject of conditionals. I’m happy to note that Field’s views and mine on the matter seem to be converging, though they certainly

[127] The previous exchanges were Field (2005), Priest (2005b), Field (2008), esp. Part V, and Priest (2010).

don't coincide yet![128]

Let us set aside for the moment the matter of restricted quantification—I'll come back to this—and just talk about the ordinary conditional.[129] My current thinking on this can be found in Priest (2009) and (2018b). In these places, I give a semantics for a relevant conditional logic of a very standard kind, using a formula-indexed binary relation. Technically the semantics endorsed in spelled out in INCL, 10.7.[130]

Field gives his preferred account in §2. It deploys a world-indexed ordering relation, rather than a sentence-indexed binary relation. But these techniques are well known to be different ways of doing much the same thing. Both of us are clear that this is the semantics for a basic logic, and might well be strengthened by constraints on the relation/ordering. Both of us admit impossible worlds. Field allows for the possibility of both moderately impossible worlds and anarchic worlds.[131] In INCL chs. 9 and 10, I give semantics with each of these—and of course, if you have anarchic worlds, some of these will behave in the same way as moderately impossible worlds. The structure of impossible worlds is an important philosophical question, though neither of us thinks that it is really important in this context.[132] So we are very

[128]Field's semantics for conditionals have gone through many iterations. (See Field (2008), (2014), (2016), (201+).) When, in the past, I have pressed him on the subject of ordinary-language conditionals, he has always said that he didn't care much about these. He is happy simply to replace natural language. I see now that he takes ordinary-language conditionals more seriously. For the evolution of my own views, see IC, ch. 6, IC2, 19.8, Priest (2009) and (2018b). I have always taken ordinary-language conditionals seriously. I have endorsed both relevant logic and conditional logic, separately—though perhaps I have never brought these two things together in the context of the paradoxes of self-reference.

[129]Field writes this as $\rhd$. I will stick to the notation of INCL, and use $>$.

[130]On the relevance (in the technical sense) of the conditional in question, see 10.11, Ex 12.

[131]In the text to fn. 11, Field says 'Priest also tends, after allowing for anarchic worlds, to ignore them, since allowing them would invalidate almost every law of conditionals. In the footnote, he refers to §§9.4.6, 9.4.7 of INCL2. A warning: INCL is a text book, and should not be taken as expressing my own views. Also, as §9.4.5 makes clear, only conditionals are at issue in this discussion of logical anarchy.

[132]I think that the importance of anarchic worlds really kicks in when one is dealing with intensional operators (TNB, ch. 1), but they also have a use in the semantics of counter-logical conditionals. Field thinks that for such conditionals one does not need all of them at once, and different contexts will determine different semantics. I think that it is simpler to have a uniform semantics, and allow context to pick out which are the worlds relevant to a conditional. (See Priest (2016b), 3.3.) These matters are of little import here.

much on the same page here.

Perhaps the main difference between us at this point is this. The frameworks in question can be used to deliver both relevant logics and irrelevant logics. Field is happy with an irrelevant logic. For example, he takes it that if B is any logical truth, so is $A > B$. Thus, since $B > B$ is a logical truth, so is $A > (B > B)$. I prefer a relevant logic. For example, the following does not seem true, let alone logically true: If every instance of the law of identity fails, then if snow is green, snow is green.) However, so far, nothing crucial seems to hang on this.[133] He also insists that the account of the conditional in "non-classical contexts"—those which contain the truth predicate—should reduce to this account in classical contexts. For me, this constraint is trivially realised, since I take the logic to be the same whether or not sentences contain the truth predicate.[134]

Let us now turn to restricted quantification. There is still a central agreement here. The problem with restricted quantification, once one foregoes the material conditional, is how to express restricted universal quantification. He and I both think that 'All As are Bs' should be understood as of the form $\forall x(A(x) \mapsto B(x))$, where $\mapsto$ is not the conditional $>$.[135] However, we give different accounts of what $\mapsto$ is. This is probably the main disagreement between us, as far as conditionals go. Field gives a somewhat complex account of the semantics of his conditional. I prefer a simpler approach.

In fact, there are several ways of doing much the same thing. In any standard relevant logic, there is a conditional operator, $\rightarrow$.[136] If the semantics is of the Routley/Meyer kind, there is a ternary relation, R, such that:

- $A \rightarrow B$ is true at world w iff for all a and b such that $Rwab$, whenever

[133] Since Field has a semantics with impossible worlds, I did wonder why he did not endorse a relevant logic, when one would come at no apparent cost. He says that the usual relevant logics are not adequate to account for the normal English conditional (§4.2). Agreed. But the semantics at issue now are the semantics of a relevant conditional logic. And an irrelevant logic seems to do no justice to such conditionals, as I have just observed.

[134] Why does Field not do the same—that is, take his generalisation of these semantics to be the correct logic right from the start? After all, the truth predicate is a part of natural language. I presume that there is a good answer, though it isn't clear to me. Perhaps it is because it would make the semantics of the natural-language conditional far too complex to be grasped by lesser mortals than logicians!

[135] Field writes the conditional as $\rightarrow$. I think that this is too confusing when relevant conditionals may come into play. So I will use a different notation.

[136] I note that the presence of this operator is not necessary for $>$ being a relevant conditional.

A is true at a, B is true at b.

There may also be a binary (heredity) relation on worlds, $\sqsubseteq$, such that for all A:

- if $x \sqsubseteq y$ then if A is true at x, A is true at y

One way of defining $\mapsto$ is by giving it the following truth conditions:[137]

- $A \mapsto B$ is true at world w iff for all a and b such that $Rwab$ and $a \sqsubseteq b$, whenever A is true at a, B is true at b.

Another is by simply defining it thus:

- $A \mapsto B := (A \wedge t) \rightarrow B$

where t is the logical constant of relevant logic, which is, intuitively, the conjunction of all truths.[138] An even simpler way is to define it thus:

- $A \mapsto B := (A \rightarrow B) \vee B$

These are not all exactly the same,[139] but they all share the crucial property, that $B \vDash A \mapsto B$.[140]

The crucial question at this point is whether $\mapsto$, as either Field or I define it, has the appropriate properties. So, what, then, are the appropriate

[137] As is done in Beall, Brady, Hazen, Priest, and Restall (2006)—hereafter BBHPR.

[138] This is done in IC2, pp. 254f.

[139] For example, the third, but not the first two, satisfies the inferential version of (2*) on Field's list: $\neg\forall x(A(x) \mapsto B(x)) \vDash \neg\forall x B(x)$. Whilst the first two, but not the third, satisfy the inferential version of (4a) on Field's list: $\forall x(A(x) \mapsto B(x)), \forall x(A(x) \mapsto C(x)) \vDash \forall x(A(x) \mapsto (B(x) \wedge C(x)))$.

[140] Actually, one might even consider the possibility of defining $\mapsto$ as $(A > B) \vee B$. But this raises novel complexities. For example, $\mapsto$ is not, then, transitive. It might be thought absurd that restricted universal quantification is not transitive. However, there are standard counter-examples to transitivity in conditional logic which can easily be modified to apply to quantification. Thus, suppose that there are just two candidates for election. Then the following inference seems to fail:

- Anyone whose competitor dies before the election will win.
- Anyone who wins will have a disappointed competitor.
- So anyone whose competitor dies before the election will have a disappointed competitor.

However, this is not the place to discuss these matters.

properties? Field (§3.1) gives a list of inferences he 'takes to be compelling'; BBHPR (§§2, 3) give an overlapping, more systematic, but incompatible list of desiderata. Now, some things really do seem to be necessary for an account of restricted universal quantification. For the rest, it seems to me, these are a legitimate matter of theoretical "give and take".[141]

To start with the former: both Field and I agree that one should have:[142]

[1] $A(a), \forall x(A(x) \mapsto B(x)) \vDash B(a)$

and

[2] $\forall x B(x) \vDash \forall x(A(x) \mapsto B(x))$

Both his account and mine validate these inferences.

Turning to the latter, it would be tedious to hammer through all the other examples from the lists, especially the more marginal ones. So let me just comment on the most significant disagreement here. This concerns negation. Field endorses (1_c), that is: $\forall x(A(x) \mapsto B(x)), \neg B(a) \vDash \neg A(a)$. BBHPR explicitly reject this (desideratum B1),[143] since, given [2], the principle delivers Explosion. This is fine for Field, for whom Explosion is valid anyway. It is not fine if one is to endorse a paraconsistent logic, as I do. Similarly, Field endorses (2_c): $\forall x \neg A(x) \vDash \forall x(A(x) \mapsto B(x))$, which delivers Explosion even faster.[144] Indeed, one really should not expect this inference in a dialetheic context. Suppose that everything satisfies $\neg A(x)$. It may yet the the case that some a is such that $A(a)$, as well; and there is absolutely no reason to suppose that $B(a)$. Field says that he can see no independent reason for

[141] Field seems to endorse the policy of "the stronger the better" (§3.1). Now, I have never been persuaded by arguments of this kind. I have heard them all too often in the defence of the material conditional. But even granting that strength is a desideratum, it has to be modulo other things, such as a solution to the paradoxes of self-reference. Thus, one can certainly strengthen the logics that both he and I favour by adding pseudo *modus ponens* $((A \wedge (A \mapsto B)) \mapsto B)$. We both reject this, for reasons connected with the Curry Paradox.

[142] In what follows I shall discuss the inferential versions of principles. I will return to the matter of the conditional versions towards the end of this section.

[143] To be precise, they reject the inference of contraposition for $\mapsto$, but the reason given applies equally to this principle.

[144] Ditto for 5^*, as Field notes. He says that he finds 5^* 'totally compelling' (§4.1). From a paraconsistent perspective, which has independent virtues—even if one is not a dialetheist—there is total uncompellingness.

giving up principles of this kind which does not ascribe to the restricted universal quantifier a modal character (§4.1). The independent reason is exactly paraconsistency/dialetheism. And to reject this as a ground in the present context is clearly to beg the question.

Two further points. Field objects (§4.1) that, in a dialetheic context, there are certain valid principles of inference about the restricted universal quantifier, and which he finds compelling, for which one cannot endorse the logical truth of the corresponding conditional. Certainly, but this seems no real problem. First, there are many valid inferences for which the corresponding conditional is not a logical truth, such as *modus ponens*—as both Field and I agree. Next, in most reasoning, it is the inference that does all the hard work, not the conditional. Thirdly, the failure of the rules of inference in question is a consequence of paraconsistency itself. And, as the saying goes, one person's *modus tollens* in another person's *modus ponens.*

Finally (§4.1), Field is worried by the fact that $\mapsto$ does not reduce to $\supset$ "in classical contexts". Never mind whether or not it *should*; mine does. A classical context—i.e., one encompassed by the semantics of classical logic—is one where there is just one world, and every sentence has the value 1 or 0, but not both (and not just one where there are no contradictions). In such a context, whichever of the three definitions of $\mapsto$ is used, $A \mapsto B$ is true at the world iff $A \supset B$ is.

14.3 Curry Paradoxes and Quasi-Naivety

Field points out (§5) that, for any conditional, there will be a corresponding Curry paradox. For the material conditional, $A \supset \perp$, is just $\neg A$. The corresponding Curry paradox is, hence, just the Liar, and so will have the same solution.—The failure of PEM (aka, $\supset$-Introduction) for Field; the failure of disjunctive syllogism (aka, Explosion) for me.—But what of the Curry paradox for other conditionals, and particularly the conditional involved in the T-Schema? Some have suggested that this ought to have the same solution as that for $\supset$. There is really no *a priori* reason why all conditionals should behave in the same way with respect to their Curry paradoxes; but Field points out that, in any case, the point has no force against his solution, which is to reject conditional-introduction for this conditional, too.

The same point does not apply to a paraconsistent solution, since this rejects conditional-introduction (that is, Contraction) for the conditional of the T-Scheme, but not for $\supset$.[145] The question then becomes whether the Liar and the Curry in question are of the same kind. This is an exceptionally vexed issue. I think that they are not, but this is not the place to go into that matter.[146]

In §6 Field turns to the contraposibality of the T-Schema. We both hold that $T\langle A\rangle \leftrightarrow A$, for the appropriate $\rightarrow$. He also holds that $\neg T\langle A\rangle \leftrightarrow \neg A$ (naivety). I do not (semi-naivety). Indeed, Field holds that for any A, A and $T\langle A\rangle$ are intersubstitutable—at least in non-intensional contexts. Thus, take a conditional that satisfies $\neg A \leftrightarrow \neg A$. Intersubstitutivity gives. $\neg T\langle A\rangle \leftrightarrow \neg A$. For me, take any dialetheia, A. Then A is true and false; so $T\langle A\rangle$ is certainly true. But, generally speaking, there is no reason why A should not be just true.

Of course, if one is a deflationist about truth, and holds that A and $T\langle A\rangle$ have exactly the same content, then intersubstitutivity follows. I have never been a deflationist about truth, however.[147] And if A and $T\langle A\rangle$ really do have the same content, it follows that if someone believes A, they believe $T\langle A\rangle$, and vice versa. But it seems that someone can believe one without the other—if, for example, they have slightly odd views about truth.

Notwithstanding, it is open to a dialetheist about the semantic paradoxes to endorse naivety. Beall, for example, does.[148] I am inclined against this. For a start, it spreads contradictions beyond necessity, turning *any* dialetheia into a dialetheia about truth.[149] Moreover, semi-naivety permits one to draw useful distinctions. Thus, one can express the thought that A is true-only by saying that $T\langle A\rangle$ and $\neg T\langle\neg A\rangle$. Given naivety, these two are equivalent (modulo double negation): there is no distinction between something's being true and its being true-only. Of course, this does not articulate the distinction in a way that enforces the consistency of truth-only. There will be sentences which are false and true only. (Such as the liar in the form $\neg T\langle A\rangle$, though perhaps not in the form $T\langle\neg A\rangle$, as Field notes.) The distinction is expressed, none the less.[150] More of this later.

A standard fixed-point construction shows that a model for a language

[145] I assume here that the conditional of the T-Schema is not $\supset$, though this is not obvious. See Priest (2017).

[146] On this, see Priest (2017), §15.

[147] See, e.g., IC, ch. 4.

[148] Beall (2009).

[149] See IC, 4.8.

[150] Again as Field notes, dual considerations apply to his approach.

without T can be extended conservatively with a naive truth predicate. Since semi-naivety is weaker than naivety, the addition of a semi-naive truth predicate is also conservative. In a previous essay[151] Field called this 'uninteresting', since it does not show how the contraposed truth predicate can fail. It was not meant to: it was simply a proof of conservative extension. But in response to Field, I gave a non-triviality proof which shows how it may do so.[152] In particular, there are A's such that $\neg A$ holds in the model, but $\neg T\langle A\rangle$ does not. The contraposed T-Scheme is therefore invalid, since it has invalid instances.

In the model, the contraposed T-Schema fails only for T-free sentences. I do not take this to show that it fails only under such conditions. That was not the point of the construction, which was just to show that one may have a model of the T-Schema in which its contraposed form is not valid. So it is a fair question to ask when one may have $\neg A$ without $\neg T\langle A\rangle$. This will happen when $A \wedge \neg A$ holds, but $T\langle A\rangle \wedge \neg T\langle A\rangle$ does not. For the second of these to be true, something (else) must force us to suppose so. (Contradictions should not be multiplied beyond necessity.) The model shows that if A is T-free nothing so forces us. Sometimes, as we have seen, we are so forced; for example, when A is $\neg T\langle A\rangle$. But there seem to be T-ful sentences where this does not appear to be the case; for example when A is $T\langle \neg A\rangle$. So when are we forced to accept the T-ful contradiction? The answer, I think, will be given by models of the T-Schema which are, in an appropriate sense, minimally inconsistent. However I have nothing useful to say on that matter at present.

Field notes that someone might object that a non-classical theory of truth is not really about *truth*, since its saves (§6):

> the truth schema in name only. The charge is that the connective $\leftrightarrow$ in the non-classical logician's preferred version of "$True(\langle A\rangle) \leftrightarrow A$" is some contrived connective, far from what motivates the idea that $True(\langle A\rangle)$ should be equivalent to A.

A naiveist may reply that the equivalence is best understood as intersubstitutivity. But a semi-naiveist may reply, instead, that $\leftrightarrow$ is not at all contrived. It is exactly what we mean when we say 'if and only if', in the context of the T-Schema. Of course, what we do mean is contentious. But to assume

[151] Field (2008), p. 371.
[152] Priest (2010), §11.

it is not what a non-classical logician says is just to beg the question against them. I note also that if the relevant biconditional is the ordinary English (bi-)conditional, >, then this does not contrapose. $True\langle A\rangle <> A$ does not entail $\neg True\langle A\rangle <> \neg A$; and we have semi-naiveism.

14.4 *L'Affaire Gödel*

In §7 Field raises matters to do with Gödel's Theorems.[153] This takes us back to where the dialetheic journey started for me.[154] The idea that the theorems might motivate dialetheism was a provocative but simple one. But, like all philosophical ideas of any interest, matters have turned out to be more complex. Let me say how things now appear to me, especially *vis à vis* Field's comments. Let us start with the relatively uncontentious matters, and work our way up to the most contentious.

Behind Gödel's proof of his first incompleteness theorem, there is an obvious paradox of self-reference. Let us write $Prov\, x$ for 'x is provable', and angle-brackets as a name-forming device. Anything provable is so. That is:

- $Prov\langle A\rangle \rightarrow A$

Let us call this, for want of a better name, *Löb's Principle*.[155]

One can, of course, ask for the justification of Löb's Principle. And if proof is proof in some particular formal systems of arithmetic, the schema is known to be unprovable, on pain of contradiction (Löb's Theorem). But here we are not yet dealing with proof in some formal system, but proof *simpliciter*—what we might call the naive notion of proof. To be provable in this sense is simply to be established as true. Löb's Principle seems, then, to be a plain *a priori* truth.

Now, consider the sentence 'this sentence is not provable'; that is, a sentence, G, of the form $\neg Prov\langle G\rangle$.[156] Substituting this in Löb's Principle gives

[153] More on this topic, see the discussion of Shapiro, §21 below.

[154] See Priest (1979) and (1984).

[155] And *Prov* means provable from things *including* Löb's Principle. So it is not provability in some other system, as Field moots in his last paragraph of the section.

[156] The construction of such a sentence in formal arithmetics requires that the primitive recursive function of diagonalisation be representable by a function symbol. If we have only the usual successor, addition, and multiplication function symbols at our disposal, we can construct only a G materially equivalent to $\neg Prov\langle G\rangle$. This fact has no material effect on the considerations to follow.

us $Prov\langle G\rangle \to \neg Prov\langle G\rangle$. Hence, we have proved $\neg Prov\langle G\rangle$. But this is just G. So we have demonstrated $Prov\langle G\rangle$. We have then established both the Gödel sentence and its negation. Let us call this *Gödel's Paradox*. The paradox is clearly a paradox very similar to the "Knower Paradox", and is in the same family as the Liar.[157] If one subscribes to a dialetheic solution to the Liar, then one should equally subscribe to a dialetheic solution to this.

It is clear that one can avoid the dialetheic conclusion if one rejects the PEM, and so the inference to $\neg Prov\langle G\rangle$, as I presume Field would. However, such a move requires a justification—and one not simply of the form 'if one does not reject this, a contradiction will arise'. Note that $Prov\langle A\rangle$ entails $T\langle A\rangle$, but not vice versa. So $Prov\langle A\rangle \vee \neg Prov\langle A\rangle$ does not entail $T\langle A\rangle \vee \neg T\langle A\rangle$. A justification is therefore required independent of the failure of the PEM for T-ful sentences. The case for the failure of the PEM for T-ful sentences is, I take it, something like Kripke's: truth and falsity are determined "from the ground up", and there is nothing to determine the truth of ungrounded sentences. Hence, they are neither truth nor false. There is, as far as I can see, no similar argument to be made for provability. So at this point, Field seems to have offered no solution to this paradox.

The notion of proof I have been talking of till now is an informal notion. Let us now turn to how matters stand if we are dealing with some formalisation of the naive notion of proof, i.e., representing it as proof within some formal axiom system. In IC, 3.5 (pp. 49f.) I gave an argument to the effect that, given an intuitively sound notion of formal proof, $Prov$, one can give an equally intuitively sound argument for $Prov\langle A\rangle \to T\langle A\rangle$. Löb's Principle of course follows from this and the T-Schema. Field has convinced me that, intuitive as this proof may be, one is not entitled to it, for reasons to do with Curry's paradox, at least if the system of proof is an axiom system which uses *modus ponens*[158]—though I do think that this sort of argument is what underlies claims that the Gödel sentence, G, for, say, Peano Arithmetic (hereafter, PA) is true. However, for the paradoxical argument to run, one does not have to bring truth into it, as I have shown above. And Löb's Principle itself strikes me as something one should have in any formal system which attempts to capture our intuitive notion of proof—just as much as one should have the T-Schema in any formal system which attempts to capture our intuitive notion of truth.

[157] See BLoT, 10.2.

[158] See Priest (2010), §5.

None of this assumes that the axiom system we are talking about is one in which the theorems are recursively enumerable (re), or even arithmetic.[159] (Though this is sometimes built into the definition of a formal system, there is no technical necessity to do this—the usual definition of a formal proof works for any set of axioms.) In this case, there is no reason to suppose that the proof predicate for the system can be defined in purely arithmetic terms.

There are, however, arguments to the effect that for a formal system adequate to our naive notion of proof (for, say, arithmetic), the theorems are re.[160] I don't claim that these are definitive, but they have a certain force. And if what they show is correct, the proof predicate for the system, and the corresponding Gödel sentence, is expressible in purely arithmetic vocabulary—where, presumably, the PEM is not at issue. It follows that the set of true purely arithmetic sentences is inconsistent. Nor is there anything technically unfeasible about this. We know that there are perfectly sensible re theories in the language of arithmetic which are complete (in the sense of containing everything true in the standard model), but inconsistent. Unsurprisingly, each validates both its Gödel sentence the negation thereof.

If the above is correct, then the set of sentences true in the language of arithmetic is inconsistent. It does not follow that PA is inconsistent; nor have I ever claimed that it is. Shapiro (2002) claims that I am, none the less, committed to this. However, his argument fails. It invokes the claim that all recursive sets/predicates are representable in PA. Now, a binary relation, Θ, is representable in a theory, $\mathfrak{T}$, if there is some formula of two free variables, $\theta(x, y)$, such that:

- if $\langle n, m\rangle \in \Theta$ then $\theta(\langle n\rangle, \langle m\rangle) \in \mathfrak{T}$
- if $\langle n, m\rangle \notin \Theta$ then $\neg\theta(\langle n\rangle, \langle m\rangle) \in \mathfrak{T}$

where $\langle k\rangle$ is the gödel number of the numeral of k. Thus, if Θ is the proof relation then, in the present scenario, it is recursive. So if PA is consistent then, clearly, Θ cannot be represented in PA—though it may be representable in an inconsistent arithmetic.[161] Of course, there is a standard proof that all decidable sets are representable in PA. One may therefore ask where that breaks down. I point this out in IC2, 17.7, where all this is discussed.

[159]See IC, 17.5.

[160]These are given in Priest (1984), §6, and IC, 3.2.

[161]I note also that this opens up intriguing new possibilities in computation theory. See, e.g., Weber (2016b).

This takes us to argument which Field's text displays in §6. To the extent that this concerns provability in PA (or similar system), then, whatever else there is to be said about things, the matter is the same: the argument simply assumes that all recursive relations can be represented in PA. They are not, if those relations are inconsistent and PA is not.

However, to the extent that the argument is taken to concern, not provability in an axiomatic arithmetic, but truth in the language of arithmetic, the matter is different. Field sketches an argument of mine, and then says (for *reductio*), that 'we might equally argue' in terms of the reasoning he then gives. Now, first, what should one make of his argument?

This takes a dialetheic sentence from outside the language of arithmetic, Q, to show that a sentence within the language is dialetheic. The sentence is $\langle Q \rangle = \langle Q \rangle$, and it is perhaps not so surprising that that sentence is dialetheic. After all, we know that in any of the inconsistent models of arithmetic there are dialethic identities. But the argument can be generalised. Let s be any non-empty decidable set of natural numbers. (In Field's case, this is $\{\langle Q \rangle\}$.) This is defined in the language of arithmetic (that is, defined in the true theory—whatever that is) by a formula $A_s(x)$. We now consider the set $s' = \{x : Q \wedge x \in S\}$.[162] Since Q, $s' = s$; and since $\neg Q$, $s = \varnothing$. Hence s' is defined in the language of arithmetic by $A_s(x)$ (and $x \neq x$; the matter there is the same). Take any $n \in s$. Then $n \in s'$, so it follows that $A_s(\mathbf{n})$ (where $\mathbf{n}$ is the numeral of n); but $n \notin s'$, so $\neg A_s(\mathbf{n})$.

The main problem with this argument concerns the claim that $s' = s$. This takes us into issues of paraconsistent set theory. It has to be shown that $\forall x(x \in s \text{ iff } x \in s')$, that is, $\forall x(x \in s \text{ iff } Q \wedge x \in s)$. What the 'iff' is here depends on how one understands parconsistent set theory. There are two main possibilities.[163] The first is to formulate set theory in an appropriate relevant logic, and to take the 'iff' to be the relevant biconditional. But in that case, the sentence fails from left to right. The other possibility is to formulate set theory in LP and take the biconditional to be its material biconditional. In this case, the sentence holds. But if $n \in s$, then since $\neg(Q \wedge n \in s)$, so does its negation. In this approach, extensionality tells us that $\forall x(x \in s \equiv x \in s') \supset s = s'$. And since this is a material conditional, we cannot detach the identity—even in a default form, since the antecedent is

[162] *Doppelgänger* of this kind are well known in paraconsistent set theory. See, e.g., Weber (2012).

[163] See IC2, ch. 18, and also the discussion of Batens, S 3.2 above.

contradictory. Hence, the argument breaks in this case also.

Finally, it is clear from these considerations that this is nothing like the argument that Field extracts from my texts, and of which has argument is supposed to be a *reductio ad absurdum*. That argument is to the effect that the naive provability predicate can be expressed in the language of arithmetic, and so that the paradoxical argument concerning it can be reproduced in the language of arithmetic. This is clearly quite different. In particular, it does not depend on constructing a *doppelgänger* set using a dialetheia.

Finally, we come to the most contentious matter. Field say that the claim that PA is inconsistent, or even just that there are true contradictory Σ_1 sentences in the language of PA 'strikes him as totally incredible'. I don't for a moment doubt Field's judgments concerning his own mental states, but I do question the rationality of his certitude.

Why might one be so certain that there are no dialetheias amongst the Σ_1 sentences of the language of arithmetic? One cannot, of course, claim that PA is consistent and that every true Σ_1 and Π_1 sentence can be proved in PA (the negation of a Σ_1 sentence being Π_1). We know from Gödel's theorem itself that we cannot prove all Π_1 statements true in the standard model in PA. This is true in spades if the correct model is one of the inconsistent arithmetics. PA is radically incomplete with respect to this.[164]

A more hopeful suggestion is to the effect that it is unclear how a true statement asserting the existence of something with a Δ_0 property could also be false. The point is made by Shapiro (2002), and is answered in IC2, 17.8. The answer is to the effect that in an inconsistent arithmetic the identity relation (which is of course Δ_0) is itself inconsistent, and this can "spread inconsistency" higher up the arithmetic hierarchy.

The grounds for for certitude about the consistency of PA strike me as equally dubious. Gödel's paradox fails to be representable in PA only by a whisker, and almost by luck. And who is to say that there are not other paradoxes of this or a similar kind lurking in the area? So what is Field's certitude based on? Hardly the fact that the axioms are self-evident. The fate of Frege's axioms taught us a lesson never to be forgotten about that. Certainly not the fact that we have a consistency proof for the axioms: the proof is in a system stronger than the axioms themselves. Perhaps that we have a very clear intuitive model of the axioms? But the consistency of the

[164] Assuming that it is consistent. And if it is not, then because it is based on classical logic, it is radically unsound!

picture is no better than the consistency of the axioms themselves. The fact that we have not found an inconsistency so far? Given the infinity of possible proofs, this is not a very good induction. And I note also that there are at least some mathematicians who take the possibility of such inconsistency very seriously; for example, the (non-Hartry) Fields Medal-winning Vladimir Voevodsky.[165]

14.5 Paradoxes of Denotation

Next (§9), we turn to a crucial matter where Field and I disagree: the paradoxes of denotation, and specifically Berry's paradox. Denotation can be defined in terms of satisfaction. So any model-theoretic construction that accommodates satisfaction accommodates denotation. Except that... the paradoxes of denotation have peculiarities all of their own.[166] One is that they use some sort of description operator essentially; and once such is in the language, the proofs of standard fixed-point constructions (such as Field's) break down.

Now, IC 1.8 formalises the argument to contradiction in Berry's paradox in a logic which does not contain the PEM. So Field's solution to the Liar paradox appears not apply to it. The formalisation in IC uses a least number operator satisfying the principle:

[Mu] $\exists x A(x) \vdash A(\mu x A(x))$

the quantifiers ranging over natural numbers, appropriate precautions being taken to prevent clash of bound variables. Note that this principle says nothing about what the denotation—if any—of '$\mu x A(x)$' is when nothing satisfies $A(x)$. Nor is there anything implausible about it, even when there are cases of denotation failure.

In his analysis of the argument, Field, taking it to use definite descriptions, objects to the inference from 'there is a n such that Fn' to 'there is a least n such that Fn' on the ground that it presupposes the PEM. However,

[165]Voevodsky (2012); and his Princeton Colleague, Edward Nelson (2015). I should make it clear that I am not endorsing the work of either of these people. I merely cite them to show that some very good mathematicians do not share Field's incredulity. And just in case anyone is tempted to misunderstand what I have said here: I am not arguing against the consistency of PA; merely against our certitude that it is consistent.

[166]See Priest (2006b).

there is no such step in the argument as formalised, so the objection is beside the point.

Of course, one might think that an *analogous* objection applies to the use of [Mu] directly. It does not. When one applies Field's argument directly to [Mu], one obtains the following. Let B be an arbitrary sentence, and let $A(x)$ be $x = 1 \vee (x = 0 \wedge B)$. Now clearly, $A(1)$, and so $\exists x A(x)$. Let τ be $\mu x A(x)$. [Mu] gives $A(\tau)$; that is, $\tau = 1 \vee (\tau = 0 \wedge B)$. This entails that $\tau = 1 \vee \tau = 0$. If $\tau = 0$ then we can rule out the first disjunct, and so B follows. If $\tau = 1$, then $\neg A(0)$, since 1 is the least n such that $A(n)$. That is, $\neg(0 = 1 \vee (0 = 0 \wedge B))$, which, given that $0 = 0$, entails $\neg B$.

But if one cannot assume the PEM, this argument fails. For that 1 is the least n such that $A(n)$ does not imply $\neg A(0)$. ($A(0)$ may be "neither true nor false".) Indeed, assuming that the extensional connectives in this context work in the standard way (say of K_3), if B is neither true nor false, so is $A(0)$. Hence, the argument for $B \vee \neg B$ begs the question.

Indeed, given an interpretation of the language of arithmetic which allows for the possibility that the PEM fails, this can be extended to an interpretation for the least number operator. '$\mu x A(x)$' denotes the least number, n, such that $A(n)$ (is true), if there is such (and whatever one wants to say about the matter if this condition fails). This verifies [Mu], and is a conservative extension, which does not, therefore, deliver the PEM.

Now, as I noted, Field's discussion is predicated on the assumption that the least number operator can be defined as a definite description: $\iota x(A(x) \wedge \forall y(y < x \rightarrow \neg A(y))$. This is not the μ-operator of the last paragraph. Of course, in a classical context, the two operators coincide. This just shows that an equivalence that works in a classical context can fail is a non-classical context of the kinds that Field and I endorse. That is a lesson which has been learned in non-classical logic many times over. Field's assumption of the equivalence would therefore seem to be vitiated by distinctions drawn in his own framework.

Finally, as Field notes, the argument for Berry's paradox can be run equally well with an indefinite description operator, ε, instead of a least number operator, μ. Such operators are well known to deliver conservative extensions of the underlying logic—whether or not the underling logic is classical. (One merely augments the semantics with a choice-function, to be employed in the denotation conditions for ε-terms.) Field avers 'I don't think it's in the least clear that there there's much cost to regarding the ε-operator ... as illegitimate'. The thought, presumably, is that the notion is incoherent in some way. But it is quite coherent, both intuitively, and on all the standard semantics. So there had better be an independent argument for this claim, or this is simple *ad hoc*ery.

14.6 Expressibility and Revenge

In §§11, 12 Field turns to the topic of expressibility and revenge. He considers a predicate, M,[167] satisfying (for me) the condition (I quote):

[M1] It should be legitimate to accept $M\langle A\rangle$ iff it is legitimate to reject $\neg A$

and (for him):

[M2] It should be legitimate to accept $\neg M\langle A\rangle$ iff it is legitimate to reject A

I am not entirely clear what 'legitimate' means here, or, for that matter, the sort of 'should' that is in question. But I don't think that that is a crucial matter at the moment.

He notes that neither of us can accept the existence of a predicate satisfying the respective conditions, or All Hell breaks loose. He says that we should each take the existence of such a predicate to be 'an illusory ideal': there is no such notion. I agree with him completely. Of course, a lot more should be said about this, if the thought is not simply to be of the disappointing kind: 'if there is such an M, I'm in trouble'. For me, at least, the existence of M is delivered by Boolean negation, †. One may define $M\langle A\rangle$ as $T\langle A\rangle \wedge \dagger F\langle A\rangle$. M can, in turn, be used to state the truth conditions for Boolean negation: $\dagger A$ is true iff $M\langle \neg A\rangle$. I have said what I have to say about Boolean negation elsewhere (DTBL, ch. 5), and there is no need to repeat it here.

Field notes that there is a way of obtaining part of what would be required by using an appropriate conditional. Thus, for me, $T\langle A\rangle \to \bot$ will do some of the job. But as he points out, for reasons to do with the Curry paradox this can be no more than partial. Indeed, asserting $T\langle A\rangle \to \bot$ will not even count as a rejection of A for all speakers. A trivialist will assert it, and reject nothing.[168] Field offers a hierarchy of predicates which approximate more and more closely the (illusory) ideal. I am not inclined to follow him down

[167] M for 'true with no monkey business'.
[168] On trivialism, see DTBL, ch. 3.

this path. (Over the years, I have developed an antipathy to hierarchies—or all kinds.) I think that there are better ways of proceeding.[169]

Both Field and I think that there are some things that are true in a mundane sense (‘Field is a person’); some things are false in a mundane sense (‘Priest is a frog’), and some other things which are neither of these (‘this sentence is not true’). The pressure for the existence of the predicate M comes from a certain take on talking about things that are or are not in the third category.

I think that there is a perfectly good way of doing this.[170] Things that are in the first category are those A such that $T\langle A\rangle \wedge \neg T\langle\neg A\rangle$; things that are in the second category are those A such that $T\langle\neg A\rangle \wedge \neg T\langle A\rangle$; things that are in the third category are those A such that $T\langle A\rangle \wedge T\langle\neg A\rangle$. Of course, these predicates do not express matters consistently. In particular, there will be some things that are in more than one category, such as the liar, $\neg T\langle L\rangle$. They express them none the less.[171]

Field says ‘the absence of general [by which he means ‘consistent’] notions of non-paradoxical truth and non-paradoxical falsehood ... makes life awkward: Priest frequently uses such notions in informal statements of his position’. It makes life awkward only for those who assume that I intend to be consistent. *Caveat emptor*.

Indeed, the demand for consistency is one that cannot be met: by me or by anyone else. There is nothing that can be asserted which forces consistency. A classical logician’s assertion of $\dagger A$ does not do this. It merely guarantees that any inconsistency collapses into triviality. (In that way, it is like an assertion of $A \to \bot$.)[172]

Note that Field cannot do the dual thing. Thus, to say that something is in the middle category cannot be expressed—even allowing for a failure of PEM—by $\neg T\langle A\rangle \wedge \neg T\langle\neg A\rangle$, since this collapses into the contradiction

[169] Indeed, it is not at all clear that the hierarchy does avoid revenge problems, as Welch (2008), (2011) has shown.

[170] For more on what follows, see IC2, 20.4.

[171] As already noted in §14.3, above.

[172] Field say ‘at first blush ... a dialetheist must assert that a sentence is non-paradoxically true in order to preclude a hearer from thinking that while he believes the sentence, he also believes ... its negation’. First blush indeed. That matter is taken care of by the Gricean conversational implicature of stating the whole relevant truth. *Exactly* the same point applies to the classical logician who asserts something. How is one supposed to know that they are not a dialetheist?—a point that I have heard Field himself make (in my defence) in seminars.

$A \wedge \neg A$. This is a difference between Field and myself—and an important one.

Using the speech act of rejection, Field can express his attitude to statements in the third class. He can reject both A and $\neg A$. The trouble is that statements prefixed by force operators do not embed in propositional contexts. Indeed, if there were a predicate, M, which applied to just those statements in the third category, contradiction would arise.[173] For we could then construct a sentence, F, of the form $\neg M\langle F\rangle$. Suppose that $M\langle F\rangle$. Then F follows; that is $\neg M\langle F\rangle$. So we appear to have proved $\neg M\langle F\rangle$, that is F, and so $M\langle F\rangle$. This argument assumes the PEM, in the form $M\langle F\rangle \vee \neg M\langle F\rangle$. Hence, this must be rejected, but then we must reject the equivalent $F \vee \neg F$. So $\neg M\langle F\rangle$, and we are back with the contradiction anyway.

At this point, one can say that M is a meaningless predicate. Indeed, elsewhere Field does say something like this: it is 'ultimately unintelligible'.[174] Yet one needs a predicate of this kind if one is going to make generalisations about our three categories, such as 'Not everything is in the first or second category'. (If $\dashv$ is the force operator of rejection, $\exists A(\dashv A \wedge \dashv \neg A)$ makes no sense.)

Indeed, to add insult to injury, Field himself seems to give us such a predicate. Come back to [M1] and [M2]. Field uses the expression 'it is legitimate to reject'. There is no suggestion that this is meaningless—on the contrary. So A's being in the third category can be characterised by the predicate 'it is legitimate to reject A and it is legitimate to reject $\neg A$'.

Field and I have a long-running dispute about whether his account of the paradoxes avoids revenge paradoxes. It still seems to me that it does not.

14.7 Other Matters

I turn now to a few miscellaneous matters.

In §8 Field takes up the topics of set theory and model theory. Most of what I have say about set theory and model theory I have said in IC2, ch 18. So only a few extra comments are required here.

[173] See Priest (2005b), p. §3.

[174] Field (2008) p. 356. In §11.2 of the present essay, Field complains that I misrepresent him in Priest (2010), p. 137, since I claim that he cannot define determinacy/indeterminacy. This is unfair here. Of course I know Field's definition of D. In the passage he cites (as I would have hoped that context makes clear) I am talking not about D, but what he calls 'super-determinacy', that is M.

I have endorsed a model-theoretic account of meaning, as opposed to an inferentialist account.[175] I don't have anything to add here to what I said about the matter in DTBL, ch. 11. I might note that if the chips ultimately fall in the other direction, I don't have a problem with this.

The importance for me of a standard model is not that it delivers a compositional account of meaning: a model-theoretic account does this anyway. A model explains why one is justified in applying the logic to the subject of the model. A standard model of set theory explains why one is so justified with respect to set theory. Without that, one has to attempt some dodge, such as the Kreisel squeeze argument.[176]

In §12 Field raises the question of an appropriate paraconsistent set theory for model theory, and whether there can be a standard model in it. I have addressed this in my reply to Batens (3.2). So I need say no more about it here.[177]

In §10 Field takes up the issue of paradoxes of validity, and what he calls epistemic paradoxes. I agree with him that there is nothing much about the former which is relevant to his views about the paradoxes versus mine. The latter is a little different.

Field formulates a number of sentences concerning belief and cognate notions that might be thought paradoxical, and refers to Caie (2012). Caie discusses how a rational and perfectly introspective agent should respond to the sentence:

- I do not believe this very sentence to be true.

and advocates a solution involving a failure of the PEM.

Field does not discuss the relation of the paradoxical sentences he cites to Caie's; nor does he say much concerning his own thoughts about a solutions to such paradoxical sentences—beyond pointing out that perfect introspectibility is implausible, and that the paradoxes 'put pressure on the coherence of the epistemic notions they employ'. So I will not discuss these examples further. Nor is this the place to discuss Caie's paradox and his detailed solution. So let me just note the following.

[175] So it's not true to say (as Field does §11) that the dialetheist has no 'general notion of validity'. Validity is preservation of truth-in-an-interpretation, for every interpretation.

[176] See Priest (2010), §10.

[177] I note the application of the Collapsing Lemma invoked to deliver models of set theory does not deliver a model of inconsistent arithmetic, but tweaking the collapse will.

There is a paradox in the family concerning the sentence:[178]

- It is not rationally permissible to accept this sentence.

A dialetheist can simply accept the contradictory conclusion. Field cannot. The paradox does require one idealising assumption, namely that: if A entails B, then if it is rationally permissible to accept A, it is rationally permissible to accept B. And one might certainly doubt this in general. Thus, someone might not realise that A entails B. However, it is hard to reject the particular instance of the principle used in the argument, since the inference from A to B in question is not only very short, but is actually presented in the paradoxical argument.

The paradox also uses the premise that:

- It is not rationally permissible to accept that (A and it is not rationally permissible to accept that A).

Again, one might contest this, but since Field has not said how he would do this, there is not much more to say about the matter. At this point, Field has offered no solution to the problem. Neither, I note, does Caie's solution of rejecting the PEM appear relevant, since this is not used in the paradoxical argument. So whatever solution to the paradox Field envisages will be quite different from his solution to the Liar paradox—unlike a dialetheist.

14.8 Wrapping Field Up

In summing up in §13, Field hangs the preference for his account of the paradoxes over mine on the matter of restricted quantification (the rest is merely aesthetic!). I have argued that that matter is not as decisive as he thinks. Much more important are the facts that his account cannot handle paradoxes such as Berry's and that it does not escape revenge paradoxes. Of course, I do not expect that Field will agree with me on these matters, and knowing him, there will be more to be said about these things!

15 Girard: Possibly Impossible

In his essay, Patrick Girard takes us into the world of dialetheic conditional and modal logic, an important area. In the first part of his paper he deploys

[178] See Priest (2010), §7 and DTBL, §6.6.

a variable accessibility relation to accommodate impossible worlds. In the second, he takes the much more daring step of a semantics which is itself dialetheic. Let me comment on these two parts separately.

To accommodate conditionals with impossible antecedents in a sensible fashion, given a worlds-semantics for these, one needs some sort of impossible worlds.[179] In standard modal logic, it is normal to think of the accessibility relation, xRy, as saying that y is possible relative to x—in whatever sense of possibility is in question. If it is not the case that xRy then, in that sense, y is impossible relative to x. This notion of impossibility can then be deployed for conditionals with impossible antecedents.

Such a notion of impossibility will do for many kinds of impossiblilty, such as physical and—contradictory obligations aside—moral. Girard wants to extend this to other kinds of impossibility. This is certainly fine for some of them. Thus, false mathematical statements are normally taken to be impossible. Yet there is nothing in the semantics of standard modal/conditional logic which requires them to be true at all worlds of an interpretation. So we can, as Girard does, take these things to fail at worlds outside the R-equivalence class of worlds containing the actual world.

Exactly the same is generally true of metaphysical impossibilities.[180] There is an issue here about identity statements, however. Many people, following Kripke, hold that true statements of identity are (metaphysically) necessarily so. Thus, it is necessarily true that Hesperus is Phospherous. Now, Girard gives the semantics for only a propositional logic, and so he does not say what semantics he is using for identity. However, if we stick with classical semantics (as Girard says that he does in the first part of his paper), since 'Hesperus is Phospherus' is true at the actual world, it will be true at all worlds, possible and (Girard's) impossible. Given this, conditionals with false identity statements may well come out with the wrong value, such as: if Hesperus were not Phospherus, modern political philosophy would be badly mistaken. (Compare: if Hesperus were not Phospherus, modern astronomy would be badly mistaken.)[181]

Perhaps most importantly, the construction cannot handle counter-logicals. Thus, as Girard points out, suppose that the liar sentence, L, is both true and

[179] See Priest (2016b), 3.3; Berto, French, Priest, and Ripley (201+), and the references therein.

[180] Assuming there to be such. For a skeptical view, see Priest (201+e).

[181] To accommodate such conditionals, the truth conditions for identity need to be non-classical at impossible worlds. See INCL, 23.6, 24.6.

false, as for dialetheism about the semantic paradoxes. Then the following is true:

[L] If $L \wedge \neg L$, then a consistent solution to the semantic paradoxes is correct

This is intuitively false, though it is true in the semantics, since the antecedent is true at no worlds.

The problem with counter-logicals does not end there, though. For example, consider the conditional:

- If intuitionist logic and philosophy are correct, then $G \vee \neg G$

where G is a statement of Goldbach's conjecture. This is presumably false. Yet, given the classical semantics for disjunction and negation, the conclusion holds in all worlds, so the conditional is true. Or again, whatever the relevant notion of conditionality at issue, let the *Law of Identity* refer to the logical truth of the schema 'If A then A'. Now consider:

- If every instance of the Law of Identity fails, then if G then G

This certainly appears false, though the consequent is true in all Girard worlds. To handle counter-logicals generally, one needs, precisely, worlds where any kind of logic can hold (or fail).

In the second part of his paper, Girard advocates moving from a classical logic to a paraconsistent logic to handle dialetheism about sentences such as [L]. Naturally, I am happy to go along with this. However he goes further, giving the semantics of the logic in a paraconsistent metatheory, indeed, a metatheory based on naive set theory.

He gives two reasons for this. One is the need to handle paradoxes about worlds that may turn up in the metatheory. I am on board with this too.[182] The second is a problem noted by Martin (2014). Suppose that we are working in the logic LP. Then in all the worlds, including the actual world, @, for any A, $\neg(A \wedge \neg A)$ holds. Hence $\Box\neg(A \wedge \neg A)$, that is $\neg\Diamond(A \wedge \neg A)$ holds at @. But for some Bs, $B \wedge \neg B$ holds at @. Hence something impossible happens at @. So @ is an impossible world.

I'm not persuaded by this, since the last step doesn't follow. What follows is simply that something contradictory holds at @. We are assuming that contradictions may hold at @. Given that the accessibility relation is

[182]See Priest (2018c).

reflexive, $\Diamond(B \wedge \neg B)$ holds at @, and so $\Diamond(B \wedge \neg B) \wedge \neg \Diamond (B \wedge \neg B)$ is just one of them; some of the contradictions that hold at @ concern what is possible itself. @ is a possible world none the less. For me, a possible world is a world where the laws of logic are the same as those at the actual world, so an impossible world is a world where they are different. (See INCL, 9.7.) @ is then possible, by definition. But even for Girard, the possible worlds are, *ex hypothesi*, those accessible from @ via R. To make @ impossible, it would have to be the case that $@R@$ does not hold (and @ would then be both possible and impossible).

To implement his semantics, Girard formulates this is a naive set theory based on a relevant logic (where the conditional does not contrapose, and so Martin's argument breaks down).[183] However, the most distinctive feature of the semantics is that it dispenses with an independent notion of falsity, and uses homophonic truth conditions for negation:

- $\neg A$ is true at a world, w (in an interpretation) iff it is not the case that A is true at w

The thought that one might do this is a very natural one, and has occurred to many people.[184] A cost is that it makes the semantics themselves inconsistent, but if its underlying logic is paraconsistent, where is the problem in this?[185]

By way of reply, note that, though a homophonic semantics is very natural, it is not mandatory. A non-homophonic semantics is perfectly appropriate for many purposes, such as relating the semantics to metaphysical concerns. Indeed, I note that Girard himself uses non-homophonic clauses for the modal operators.[186]

Next, note that the homophonic truth conditions spread contradictions. Thus, any contradiction, $A \wedge \neg A$, *in* a world of an interpretation delivers a contradiction *about* it: A is and is not true at the world. Contradictions

[183] I prefer to formulate the semantics in naive set theory in a rather different way. See my reply to Batens, §3.2 above.

[184] And I note that this move is quite independent Girard's strategy concerning impossible worlds.

[185] Girard says (§3) that his semantics is not recursive, and the clauses must be thought of simply as semantic axioms. This seems to me to be a mistake. The clauses are perfectly well-grounded, and if the set theory has the means to turn a recursive definition into an explicit definition, the axiomatic clauses can be turned into such a definition.

[186] A homophonic clause for $\Box$ would be: $\Box A$ is true at a world w (in interpretation) iff $\Box$(A is true in at w).

should not be multiplied beyond necessity (IC, 4.9). Moreover, the contradictions appear to spread where they are really not wanted. PEM ($A \vee \neg A$) is a logically valid schema. But take a world of an interpretation, w, where for some B, $B \wedge \neg B$ holds. Since $\neg B$ is true at w, B is not true at w. And since B is true at w, $\neg\neg B$ is true at w, so it is not the case that $\neg B$ is true at w. So $B \vee \neg B$ is not true at w. Hence PEM is not a logically valid (as well). More generally, let w be the trivial world. Then for any A, $\neg A$ is true at w, so A is not true at w. Hence there are no logical truths.[187]

I note, finally, that moving to a paraconsistent logic certainly removes the problem concerning the conditional [L]. However, it does nothing to rectify the shortcomings of the earlier semantics with respect to counterlogical conditionals in general.[188] Thus, since $G \vee \neg G$ is true in every world, the conditional:

- if intuitionist logic and philosophy are correct, then $G \vee \neg G$

is still true.

Hence, I am not inclined to go down the homophonic path along which Girard beckons.

One final and more tangential comment on Girard's paper. He endorses the principle POS: if the antecedent of a true conditional is possible, so is the consequent. I did think this true at one time, but Dave Ripley and Yale Weiss persuaded me that I was mistaken. When evaluating a conditional, one looks at worlds where the antecedent is true, and where certain information bleeds across from the world of evaluation.[189] Normally, if the antecedent of the conditional is possible, the bleeding will not take us to an impossible world. Thus, to evaluate the conditional, 'if I jump out of a 17th floor the window, I will get hurt', the information that bleeds across concerns the laws of gravity, human biology, etc; and we evaluate the consequent at worlds where these and the antecedent hold. Nothing forces such worlds to be impossible. However, suppose that we have a correctly programmed computer that searches for a proof in Peano Arithmetic of its Gödel sentence; a light will go on iff it finds it. Now, consider the conditional: if the light were to go on, something impossible would have happened. One can hear

[187] Weber, Badia, and Girard (himself) (2016), point out a number of the fraught model-theoretic consequences of what amount to homophonic truth conditions for negation.

[188] Or conditionals with false identity statements as antecedents, unless one modifies the usual semantics for identity.

[189] See Priest (2018b).

this as true, and there is nothing impossible about the antecedent. Yet if one hears it as true, the information that bleeds across from the base-world includes the claim that the machine is working correctly; and if this and the antecedent are realised, we must be in an impossible world.[190] Of course, without this piece of information bleeding across, it would be more plausible to accept the truth of the following: if the light were to go on, the machine would not be working correctly. In general, what information bleeds across in the evaluation of a conditional is contextually dependent. POS is not, then, an appropriate *logical* constraint, though it may well hold for many (most?, normal?) contexts.

16 Humberstone: Everything you Wanted to Know about *LP* Negation...

Let us now turn to Lloyd Humberstone's paper. Humberstone and I have known each other for many years, and our intellectual paths have often crossed—over the last 18 years at meetings of the Melbourne Logic Group, but many times before that as well. His logical interests and mine certainly do not coincide, but there is a significant overlap, and I have often benefitted from his insights. The present paper shows the scholarly knowledge, logical acumen, ingenuity, and thoughtful care which characterise his work, such as his magisterial book on the logical connectives.[191]

The present paper is rich in insight, and one thing it shows clearly is that the characterisation of a number of notions is a very sensitive matter. Humberstone is clear about distinctions which many others slide over. In what follows I will restrict myself largely to addressing two issues concerning *LP* which the paper raises: the uniqueness of the characterisation of negation; and negation as a contradictory forming-operator. I will deal with these in the next two sections. In a third, I will comment on some specific remarks by Humberstone, which I was unable to integrate into the more general discussion without disrupting it too much.

[190]This particular example is due to Vander Laan (2004).

[191]Humberstone (2011).

16.1 LP Negation: Uniqueness

I take it that the fundamental property of negation is that it toggles between truth and falsity. That this holds for negation can be agreed by partisans of many different logics. I am, of course, aware that there are those who would demur: there are, after all, many theories of negation. Still, the aim here is not to defend a theory of negation, but to discuss its characteristics. Given this property, we have:

- $\neg A \in T$ iff $A \in F$
- $\neg A \in F$ iff $A \in T$

where T is the set of truths, and F is the set of falsehoods. If one thinks that $\neg A$ is true iff A is not true, then one can simply give truth conditions for negation, and the falsity conditions will go along for the ride. However, once one is contemplating logical gaps or gluts, truth and falsity have to be treated even-handedly.

Suppose that we generalise these conditions from truth *simpliciter* to truth in an interpretation. If the logic is one in which truth and falsity are the only semantic considerations then, in every interpretation:

- $\neg A$ is true iff A is false
- $\neg A$ is false iff A is true

These conditions are those of many logics. If there are no further constraints on truth an falsity, and conjunction and disjunction behave in a standard way, we have the logic FDE. If one insists that nothing can be both true and false (in an interpretation), we have K_3. If one insists that everything is either true or false, we have LP. If one insists on both, we have classical logic. (I assume that validity is defined as truth preservation in all interpretations.) Which, if any, of these constraints is correct is a separate semantic/metaphysical issue. I endorse those of LP, though this is not the place to go into the matter.[192] As I have just said, the point here is not to defend a theory but to explore its properties.

These conditions then define $\neg$ uniquely, in at least a couple of different senses. Thus, suppose there is another operator, $\neg'$, which satisfies the same conditions. That is, in any interpretation:

[192] That position is defended in IC, ch. 4.

- $\neg' A$ is true iff A is false
- $\neg' A$ is false iff A is true

then in any interpretation:

- $\neg A$ is true iff $\neg' A$ is true
- $\neg A$ is false iff $\neg' A$ is false

It follows that:

- $\neg A \dashv\vDash \neg' A$

There is also a stronger sense of uniqueness. In classical logic, an operator is called a truth function if whenever the inputs have the same truth value, the output has the same value. In the present context, we have to say, instead, that whenever the inputs have the same truth and falsity values, the output has the same truth and falsity values. In the language of LP, all the connectives (including $\neg'$ if this is added) have this property. And provided all the connective have this property, we have:

- $C(\neg A) \dashv\vDash C(\neg' A)$

If our logics are extended to ones with a world semantics, matters are slightly more complicated. Let us say that a logical operator is *truth functional in a modal sense* if whenever all the inputs have the same truth/falsity value at every world, so does the output. This condition is satisfied by the usual extension of the language of LP by modal operators, and/or a conditional connective.[193] Then the previous two bullet points still hold.[194]

Nothing in this section gainsays any of Humberstone's results. As he makes clear, e.g., in his discussion of sequents in §5, he is concerned with a notion of unique characterisation in proof-theoretic terms. By contrast, here I have given a notion of unique characterisation in semantic terms.[195]

---

[193] See INCL, chs. 9, 10, and 11a.4.

[194] There are semantics with impossible worlds in which (some) sentences are assigned arbitrary truth/falsity values at these. If the truth/falsity values of $\neg$ and $\neg'$ can come apart at these worlds, then one will still have $\neg A \dashv\vDash \neg' A$, validity being defined as truth-preservation at all possible worlds. However, one may no longer have $C(\neg A) \dashv\vDash C(\neg' A)$. The same is true, of course, of $\wedge$ and $\wedge'$, $\vee$ and $\vee'$, etc.

[195] In particular, Humberstone's function which is the same as the $\neg$ of LP, except that it maps b to f does not satisfy the condition that if A is false (i.e., b or f), $\neg A$ is true (i.e, b or t).

16.2 LP Negation: Contradiction Formation

Let us now turn to the question of whether the negation of LP is a contradiction-forming operator—hereafter cdfo.[196] First, some context. I have claimed that negation is a cdfo, and have argued that the negations of many logics do not satisfy this condition.[197] Slater agreed that negation is a cdfo, but charged that the negation of LP is itself not a cdfo.[198]

To look at the matter, we first need to address the question of what contradictories are. Traditional logic is pretty clear on this matter. Consider the following conditions. Let us call these the *naive* conditions:

- it can't be the case that both A and B
- it must be the case that (at least) one of A and B

If A and B satisfy the first but not the second, they are contraries. If they satisfy the second but not the first, they are subcontraries. If they satisfy both, they are contradictories. Note that this is what Humberstone terms 'mere' contraries/subcontraries.

But how should one understand the modal terms deployed here? Any way of making matters precise will deliver an understanding of what it is to be a contrary/subcontrary/contradictory. And as Humberstone's paper makes abundantly clear, there may be many ways one might do this, delivering different non-equivalent notions.

One way to understand these two conditions (respectively) is as follows:

- there is no interpretation in which both A and B are true
- in any interpretation at least one of A and B is true

This is pretty much what Slater means in his critique of paraconsistent negation. However, it is not a happy way of cashing out an understanding of the notions, at least if that understanding is the standard one in the history of Western logic. For a start, the notion of an interpretation is a creature of the last 100 or so years, and is alien to the thought of Aristotle, Buridan, Leibniz, Kant, etc. Worse, a standard example of contraries is: 'x is red' and 'x is green'. In modern logic, one can but write these as Rx and Gx. But

[196] Since Humberstone objects to 'cfo'.
[197] DTBL, ch. 4, and Priest (2007).
[198] Slater (1995) and (2007).

there are interpretations where something can satisfy both these predicates. So they are not contraries on this account. Similarly, a standard example of subcontraries is: 'part of x is the Northern Hemisphere', and 'part of x is in the Southern Hemisphere', as applied to geographical features on the Earth. Writing these as Nx and Sx, respectively, there are interpretations where nothing satisfies both, and so they are not subcontraries according to this account.[199]

A better understanding of these conditions is to take the modal operators deployed at face value. The conditions then become:

Con $\neg \Diamond (A \wedge B)$, i.e., $\Box\neg(A \wedge B)$

SubCon $\Box(A \vee B)$

And what is the negation sign here? It is whatever those who deployed the informal locutions intended. That is, it is the negation of ordinary language. How that behaves is, of course, a contentious matter. But those of an LP persuasion, such as myself, will take this to be LP negation.

For negation to be a cdfo in this sense, we then have:

- $\Box\neg(A \wedge \neg A)$
- $\Box(A \vee \neg A)$

And since we are dealing with negation, which is a matter of logic, and not red/green, north/south, we should expect these to be logical truths. Moreover, if our modal operator satisfies standard conditions, and in particular Necessitation (if $\vDash A$ then $\vDash \Box A$) and veridicality ($\Box A \vDash A$) then these are equivalent to:

- $\vDash \neg(A \wedge \neg A)$
- $\vDash A \vee \neg A$

[199]Another possible understanding of what it is for A and B to be contraries is mooted by Humberstone (§3): $A, B \vdash$. (Similarly, with $\vdash A, B$ for subcontraries, and with both for contradictories.) Clearly, this will not do in general, for similar reasons. One would not expect it to be a fact of logic that: $Rx, Gx \vdash$. Nor is it appropriate when B is $\neg A$. For this simply builds Explosion into the definition of a cdfo, and so begs the question agains a paraconsistentist, as Humberstone notes later in the section.

Now neither nor FDE nor K_3 satisfies both of these conditions; but LP does. So in this sense, the negation of LP is a cdfo.[200]

There other logics where these conditions are not met. Thus in intuitionist logic one has the first, but not the second; and in da Costa's C-systems one has the second but not the first. Now, the whole issue about cdfos arose originally because Richard Routley (as he then was) and I argued on the basis of this observation that the negation of the C-systems is not a cdfo, but a (mere) subcontrary-forming operator.[201] This point, then, stands.

Of course, we may, in the end, have to conclude that, according to our best theory of negation, negation is not a cdfo in this sense. But the fact that an account of negation does not deliver an operator which is a cdfo in the naive sense, is a black mark against it. So much of the history of logic would, then, have to be written off as mistaken.[202]

Let us now turn to the matter of the uniqueness of contradictories. DTBL, p. 78, says that contradictories, unlike contraries and subcontraries, are unique. How may one best understand this claim?

As I have suggested, A and B are contradictories iff $\Box(A \vee B)$ and $\Box\neg(A \wedge B)$. Now, if $\bot$ is a logical constant such that $\bot \vDash C$, for all C. Then in classical logic B is logically equivalent to $B \vee \bot$. So if A and B are contradictories, so are A and $B \vee \bot$. Hence, different formulas can be contradictories of A. Moreover in LP, B is also logically equivalent to $B \vee \bot$. So the same point applies.

However, in classical logic:

(*) $A \vee B, \neg(A \wedge B), A \vee C, \neg(A \wedge C) \vDash C \equiv B$

Hence, $\Box(A \vee B), \Box\neg(A \wedge B), \Box(A \vee C), \Box\neg(A \wedge C) \vDash \Box(C \equiv B)$. So contradictories are unique up to necessary material equivalence. One might think

---

[200]Humberstone (§1) quotes a passage from DTBL and says, 'I read this as saying that for a one place connective, #, to count as a c[d]fo in logic S it is necessary and sufficient that we have, for an arbitrary formula, A, of the language of that logic ... $\vdash_S A \vee \#A$ and $\vdash_S \#(A \wedge \#A)$', and points out that these conditions are satisfied when $\neg A$ is $\top$. Whether or not this is a reasonable interpretation of the passage, it is not what I intended, as I hope is now clear. That the wide-scope # is negation is presupposed.

Yet another way of interpreting negation as a cdfo is Wansing's as discussed by Humberstone in his Appendix B. This is that $A \vdash \neg\neg A$ and $\vdash \neg(A \wedge \neg\neg A)$. These conditions are satisfied in intuitionist logic. The failure of Double Negation in intuitionist logic makes these a particularly unsuitable way of cashing out the naive (and traditional) understanding of the notion.

[201]Priest and Routley (1989a), p. 165.

[202]For further discussion, see DTBL, esp. 4.4.

of $\Box(B \equiv C)$ as saying that B and C express the same proposition. In this case, there is a unique proposition which is the contradictory of A.

This is not true in *LP*, where (*) does not hold. (Make A both true and false.) However, we do have:

(**) $A \vee B, \neg(A \wedge B), A \vee C, \neg(A \wedge C) \vDash (C \equiv B) \vee A!$

where $A!$ is $A \wedge \neg A$. Hence, we have $\Box(A \vee B), \Box\neg(A \wedge B), \Box(A \vee C), \Box\neg(A \wedge C) \vDash \Box((C \equiv B) \vee A!)$. Let us say that B and C are *materially equivalent relative to* A if $(C \equiv B) \vee A!$ Contradictories of A are then unique up to necessary material equivalence relative A. If one thinks of $\Box((B \equiv C) \vee A!)$ as saying that B and C express the same proposition relative to A, then the contradictory of A is a unique proposition relative to A.

Note that this is not the same with respect to subcontraries, mere or not (or contraries—the dual considerations apply). Classically:

- $\Box(A \vee B), \Box(A \vee C) \nvDash \Box(B \equiv C)$

So subcontraries of A are not unique up to necessary material equivalence. And in *LP*:

- $\Box(A \vee B), \Box(A \vee C) \nvDash \Box((B \equiv C) \vee A!)$

Hence, subcontraries of A are not unique up to necessary material equivalence relative A.

16.3 Some Further Comments

In this final section, I will comment on a number of further claims in Humberstone's essay, which I think merit a mention.

(i) In §2 he quotes a passage from DTBL and says that it is rather hard to understand. Let me see if I can do better. Let us suppose that some things are neither true nor false; that is, for some As neither A nor $\neg A$. One might object that $\neg$ is not really a cdfo, since the PEM is violated. (So $\neg A$ should hold in all the cases that A *fails*.) Whether or not this is a good objection, the parallel objection to someone who holds that some things are both true and false—that is, for some As, both A and $\neg A$—does not work, simply because logical gluts do not violate the PNC, as we have seen.

(ii) In a footnote to the same passage, Humberstone says that he does not understand the notion of surplus content. I agree that its explanation in Priest (2007) is not very clear. What I had in mind was this. Classically, $\neg(A \wedge \neg A)$ rules out $A \wedge \neg A$. In a paraconsistent context, it does not, so there can be a "surplus".[203] Of course, classically a contradiction entails everything, and paraconsistently, it does not. So in that sense, it is a classical contradiction that has a surplus.

(iii) A little later in the the section, Humberstone notes that defending the view that $\neg \Diamond (A \wedge \neg A)$ is a statement of the PNC by pointing out that it allows for $\Diamond(A \wedge \neg A)$, is, in effect, question-begging against Slater. But of course, to say that it does not do so because of this, is equally question-begging against me. Questions of onus of proof are tricky. However, in this context, it was Slater who initiated the argument by claiming that the negation of LP is not a cdfo; and so the onus of proof must fall on him.

(iv) In fn 22, Humberstone says that I follow the motto: when the going gets tough, go homophonic. I don't think that's right, if by 'homophonic' one means giving the truth conditions of some notion using that very notion. Though this may be a perfectly sensible strategy, it is not one that I fall back on when 'the going gets tough'. He references a passage in DTBL, but the point there is not about homophony but about using the same logic in the object and metalanguage—a quite different point, and one which I do not fall back on: it's right up front.

(v) At the end of §3, Humberstone comments on a passage in DTBL, where I argue that Con and Subcon deliver Double Negation. He points out, correctly, that this presupposes the uniqueness of contradictories, and objects that Con and Subcon do not deliver such, on the ground that they hold when negation is interpreted as $\top$. As I intend them, however, Con and Subcon require that we are dealing with the correct negation—that of LP—as I have already said. It may fairly be replied that since Double Negation holds in LP, the argument does, in a certain sense, presuppose this.

(vi) In a digression at the start of §6 Humberstone comments on a passage in DTBL where I discuss a certain notion of logical pluralism. I claim that classical negation and intuitionistic negation have different meanings,[204] but that one cannot put them together in a single language. He says that he infers from the latter fact that they do not have different meanings. I infer that it may not be possible to combine independently meaningful things. After

[203] I note also, as, in effect, does Humberstone in §4, that $A \wedge \neg A$ entails $\neg A \wedge \neg\neg A$. So, in this sense, a dialetheia is neither true nor false. Even more surplus!

[204] Incidentally, in a footnote, Humberstone asks what truth conditions I had in mind for intuitionistic negation. It was those in Kripke semantics.

all Boolean negation—assuming it to be meaningful—cannot be combined with the intuitionistic conditional without this collapsing into the classical conditional; similarly, Boolean negation cannot be combined with a naive truth predicate without total collapse.[205] Again, as they say, one person's *modus ponens* in another's *modus tollens*.

17 Istre and McKubre-Jordens: Relevant Conditionals and Naive Sets

With Erik Istre and Maarten McKubre-Jordens we venture into the world of relevant naive set theory.

I have always thought that the correct approach to the paradoxes of set theory is to accept the naive principle of set abstraction and the consequent paradoxical contradictions, but to use a paraconsistent logic to prevent the spread of contradiction.[206] A natural thought is to use an appropriate relevant logic as the paraconsistent logic in question. In the late 1970s Brady's proof showed that this could be done without triviality.[207]

Of course, showing what can't be proved is not sufficient. One also needs to show what *can* be proved. In particular, one needs to show how one can carry out natural set theoretic reasoning, such as that pursued by Cantor. The aim was never to be revisionary, in the way that intuitionism is. And this is not a trivial problem. The set theoretic axioms are strong, but to avoid triviality the underlying logic must be much weaker than "classical" logic. Simple reasoning about basic set theoretic operations, such as unions and compliments, is routine,[208] but after that, the going gets tough.[209]

This problem exercised many of those in the Canberra logic group in the late 70s and early 80s, including Routley, Meyer, Brady, Mortensen, Slaney, and myself. The paper by Istre and McKubre Jordens (hereafter I&MJ) well explains the sort of problems we hit. I am not persuaded by all the ones they note. Thus, they claim (§3) that there is a problem with meta-rules, since these 'make a claim about what proofs can be constructed'. Not so; they

[205] See DTBL, 5.5.
[206] See IC, ch. 3.
[207] Brady (1989).
[208] This was all spelled out in Routley (1977).
[209] And without the PEM one seems completely hamstrung, as IMJ note (fn 1).

are constitutive of what *counts* as a proof.[210] However, problems of the kind they document are very real. So much so that, at the time, all of us, I think, gave up trying to solve them. I certainly did. I always hoped that someone might be able to do it; but after that I started to think about other ways of achieving the "classical recapture" in set theory, and especially how to use a model-theoretic approach to achieve this end.[211]

A highly significant advance on the problem was made by Weber in his Melbourne PhD thesis of 2009, where he showed how to obtain most of the standard results concerning transfinite ordinals and cardinals.[212] His trick was not to try to reconstruct the classical proofs—as we had been trying to do—but to develop quite new proofs which deploy inconsistency essentially. (And whether or not one is persuaded by the legitimacy of Weber's techniques, they are certainly fascinating.)

That work, however, still leaves much to be done. In particular, one needs to show that things like model theory can be done in this relevant setting. Weber has continued to work on this, sometimes tweaking the underlying relevant logic in the process;[213] and maybe something like this can be done. But at the moment, problems of the kind that I&MJ document still stand in the way. Without solutions to these, I think that the best approach to paraconsistent set theory is the alternative one I indicated above.

18 Mares: They are, are They?

For a long time, relevant logic was something of an outlier in the family of non-classical logic, and something of an ugly duckling. That is started life simply in axiomatic form didn't help. And when a robust semantics appeared, in the shape of Routley/Meyer semantics, matters seemed to become worse. First, the semantics appeared complex and somewhat awkward. Next, it became clear that the family of relevant logics was enormous. That, in itself, is not a serious problem. The same is true, of course, of modal logics. There, one just has to say which notion of modality one has in mind, and choose the most appropriate system for it. The same for relevant logic: one

[210] A very minor comment on their paper. They say in §2 that $A \rightarrow (A \wedge A)$ is not provable in DKQ. This is surely a slip. It follows from their A1 and A6.

[211] See IC2, ch. 18, Priest (2017), §§10, 11, and 3.2 above.

[212] See Weber (2010), (2012).

[213] See, e.g., Weber (2016a).

just has to say which notion of conditionality one has in mind, and choose the most appropriate system for that. The problem is that conditionality is a much more vexed issue than modality, and so it is much less clear how to choose. And the semantics were, in a sense, too good: one could do nearly anything with them! Not only that, but the machinery deployed, notably the ternary R and the Routley $*$-operator seemed to want for a plausible philosophical interpretation. Hence their properties could provide no guidance in this matter.

The situation is now much improved. Philosophical issues are slowly coming into focus.[214] The formal machinery has also been shown to have interesting technical properties.[215] Moreover, intimate connections have emerged with the family of sub-structural logics, with information and its flow, and with other accounts of conditionality.[216]

It can now be said that, at least as far as propositional logic goes, the machinery is well under control, both philosophically and technically. The same is not true of its extension to a first-order context, however. Questions concerning predication, quantification, identity, are still vexed. Ed Mares, another stalwart of the band of Australasian logicians, has been at the forefront of work in this area, trying to bring order into the complexity.[217] His paper here continues his investigations of identity in the context of relevant logic. Drawing on an analogy between identity, as a binary relation between objects, and biconditionality, as a connective between sentences, he provides two semantics for identity, and discusses their properties.

As he says, there is a clear analogy between identity and the biconditional. It certainly does not mandate the view that the two behave in an isomorphic fashion, and, as Mares notes, analogous principles may well be semantically independent.[218] However, the analogy is, at the very least, highly suggestive. Mares' first semantics is simple and clean;[219] the second, somewhat less so. In part, this is due to the presence of the ternary R; but only in part. The

[214] See, e.g., Restall (1999), Beall *et al* (2012).

[215] See, e.g., Urquhardt (1984).

[216] On which see, respectively, Restall (2000), Mares (1996), and Beall *et al* (2012) again.

[217] As well as in investigations of the philosophical basis of relevant logic, as in Mares (2004).

[218] In particular, there might well be good reasons to endorse what Mares calls FullSub, but not Pseudo *Modus Ponens*.

[219] A small comment. He says (§4) that the normal worlds of N_* are complete. This is not so. See INCL 9.6.

semantics has a feeling, often voiced concerning the original Routley/Meyer semantics, that they are being "rigged to get what you want". The feeling is exacerbated by the presence of the predicate Ind. Given the need for this, one will naturally ask: why not give the truth conditions of identity using this, and have done with it?

Still, the semantics are clever and interesting. Rather than discuss the details further, I will just comment on one of the more general matters Mares raises. That is, the schema he calls *FullSub*:

- $(a = b \wedge A_x(a)) \to A_x(b)$

This is a very natural principle concerning identity; and one which, for the appropriate As, one would need good reason to hold to be incorrect.

Of course, one should not expect this principle—or even more restricted versions of substitutivty—when A may contain intensional verbs, such as 'believes', 'admires', etc. One can believe that Richard Routley was Richard Routley without believing that Richard Routley was Richard Sylvan. Or one can admire Routley without admiring Sylvan, believing this to be a different person. But there appear to be counterexamples even in the more limited context of the vocabulary of first order logic. In particular, substitution into the scope of a conditional appears to have counter-examples. Thus,[220] it is clearly true that:

- If the Morning Star is (were) not Venus, modern astronomy is (would be) badly mistaken.

But the Morning Star is Venus. However, it does not follow that:

- If Venus is (were) not Venus, modern astronomy is (would be) badly mistaken

It would not be modern astronomy that is mistaken, but modern logic. Or again, consider:

- If I am (were) not Graham Priest then, I am not (would not be) the author of *In Contradiction*.

But I am Graham Priest. Yet is does not follow that:

[220] See INCL, 19.5.

- If Graham Priest is (were) not Graham Priest then, he is not (would not be) the author of *In Contradiction*.

Whatever the consequences of the failure of the Principle of Identity are, the consequent doesn't seem to be one of them.[221]

Next, in the semantics of INCL2, 24.6 (though not in Mares' semantics) it is the Subset Constraint that validates FullSub. The constraint is to the effect that the extension of = at a non-normal world is a subset of its extension at normal worlds (the set of pairs $\langle d, d\rangle$, where d is a member of the domain). Concerning this constraint, Mares says (§8):

> There is a problem with the subset constraint. Consider a model in which 'Hesperus is Mercury', does not hold in any normal situation. Then, in no non-normal point is Hesperus identical to Mercury. So we have valid on the model
>
> > Hesperus is Mercury implies that the moon is made of green cheese.

Let us disentangle what is going on here.[222]

For a start, the subset constraint does not give rise to "fallacies of relevance", at least in one sense. It can be shown, for example, that in the quantified relevant logic B, with constant domain, and identity satisfying the subset constraint, for any logical truth of the form $A \to B$, A and B have a predicate (maybe identity) in common.[223] In particular, then, with obvious notation $h = m \to M$ is not a logical truth.

Moreover, it is not a problem that this conditional is true in some models After all, one can make any non-logical truth hold in some model—e.g., 'Hesperus is Mercury' and 'Some red things are not coloured'. Irrelevant conditionals could simply be like this.

However, there is a genuine worry here. Consider the model in which truth values get assigned correctly. Then, assuming the necessity of true

[221]Substitutivity into a conditional holds in Mares' systems. However, the conditionals in these systems are essentially entailment conditionals, and the conditionals in the counterexamples are hardly of this kind. One might not, therefore, be too worried by them.

[222]I note that Mares uses the word 'implies' for the conditional. I think it best to avoid this. In my experience it is at the root of much confusion in students between the conditional and the validity relation.

[223]The proof extends that of INCL, p. 220, ex. 11. For details, see Priest (200+f) §3.

identities, in that model $h = m$ does not hold at any normal (possible) world, and so, given the Subset Constraint, at any world. It follows that $h = m \to M$ is true at the base world of the model, and so true *simpliciter*. This certainly does not seem to be right.

One can, of course, have a restricted form of FullSub, where one does not substitute into a conditional context (and more generally, a context whose evaluation requires a world-shift). To do this, one requires a contingent identity logic (as noted in INCL, 19.5.8). Mares indicates that this is the way he is inclined to go. He motivates the position as follows. Suppose that in our models true identities are necessarily so. That is, if an identity statement is true at the base world, it is true in all normal (possible) worlds. And suppose that 'Hesperus in Venus' is thus true. Then (§8):

> the following is valid on the frame:
>
> > Hesperus is Mercury implies that Mercury is Venus.
>
> This sentence is, I think, true... But I do not think that the counterfactual version is true:
>
> > If Hesperus were Mercury, then Mercury would be Venus.
>
> In stating the counterfactual one is asking the audience to imagine what the world would be like like if Hesperus were not Mercury. They might imagine that a bright star someone pointed to at in the evening turned out to be the closest planet to the sun, and that there is another planet between the Earth and the sun which is Venus.

Now, I am not persuaded that there is a difference of kind between indicative and subjunctive conditionals. In particular, there does not seem to me to be much difference between conditionals with antecedents in the present indicative and the present subjunctive.[224] Moreover, someone who holds that conditionals with necessarily false antecedents are vacuously true[225] might well say that, in the scenario envisaged by Mares, a person is not imagining that Hesperus, *that very planet*, is not Mercury; merely that the name 'Hesperus' had been pinned on some other planet.

[224] For further discussion, see Priest (2018b).
[225] Such as Williamson (201+).

Yet I certainly agree that in evaluating conditions such as the one in question, one does need to consider worlds where Hesperus (that very planet) is not Venus (that very planet).[226] This may be done by deploying a contingent identity semantics.[227] Objects have parts, avatars, or whatever,[228] which may vary from world to world, and the truth conditions of predicates (including the identity predicate) make reference to these. This does not require true identities to fail at normal worlds (though it is compatible with this).[229] The worlds where true identities fail may just be a impossible worlds. As long as there are some worlds where they fail, FullSub will hold when substituting into non-conditional contexts, but fail otherwise. Indeed, substitutivity in the even stronger[230] form $a = b, A_x(a) \vDash A_x(b)$ will fail if substituting into a conditional context.[231]

19 Read: Bradwardine Comes Back from the Dead

Stephen Read is my oldest philosophical colleague, friend, and coauthor. We have shared thoughts and debated ideas, both in my two periods in St Andrews (1974-1976, 2000-2013), as well as between those times and after them. I have learned much from him in the process. We share many interests, most prominently logical paradoxes, relevant logic, and medieval logic—though I, unlike Read, am very much an amateur in the last of these. Indeed, I owe my interest in it entirely to him. In my first year in St Andrews, he opened my eyes to the richness of logic in the period.

One of Read's very significant achievement over recent years is to have resuscitated the solution to paradoxes in the family of the liar advanced by Thomas Bradwardine (1300-1349), articulating and defending it with all the resources of modern logic.[232] Behind this solution is a theory of truth and

[226] See Berto, French, Priest, and Ripley (201+).

[227] See INCL, ch. 17.

[228] Mares pointed out to me that the notion of an avatar is in some ways similar to Castañeda's (1989) notion of a guise.

[229] See TNB, 2.9.

[230] Or weaker, depending on which way you think is up!

[231] INCL, 24.7.10.

[232] His many papers on the topic are referenced in his paper here, but not a bad place to start is with Spade and Read (2017).

signification. A sentence may signify many things. Indeed, according to Bradwardine it signifies everything that follows from it. A sentence is true if *everything* that it signifies is the case. So using $\mathbf{Sig}(s, p)$ to express that s signifies that p:[233]

- $T\langle A\rangle \leftrightarrow \forall p(\mathbf{Sig}(\langle A\rangle, p) \rightarrow p)$

Given that $\mathbf{Sig}(\langle A\rangle, A)$, the left-to-right direction of the T-Schema is forthcoming. But given that A will signify other things as well, the right-to-left direction is not, so the paradoxical argument is blocked. Bradwardine argues that the liar sentence ('this sentence is false') signifies not only that it is false but that it is true. Hence, signifying a contradiction, it is false. So something it signifies (viz., that it is true) is not the case. Since the right hand side of the biconditional is false, one cannot infer the left.

In 'Read on Bradwardine on the Liar' (hereafter RBL)[234] I raised a couple of problems for the Bradwardine/Read account. One was that the account seemed unable to account for paradoxes of denotation, such as Berry's paradox. In his paper here, Read cleverly takes up the challenge to show that it can.[235]

The paradoxes of denotation require descriptive terms of some kind. For present purposes, and since we are dealing principally with Berry's paradox, which concerns natural numbers, let us suppose that these are formed with the least number operator, μ (*the least number such that*). (Read uses a definite description operator, ι. Nothing hangs on this fact: exactly the same considerations apply to each.) The argument for Berry's paradox[236] then depends on two standard principles concerning such terms and their denotations:

- $D(\langle t\rangle, x) \leftrightarrow t = x$
- $\exists xA \rightarrow A_x(\mu xA)$

[233] One might also add the clause $\exists p\mathbf{Sig}(\langle A\rangle, p)$ to the right hand side, but since $\mathbf{Sig}(\langle A\rangle, A)$, this is redundant.

[234] Priest (2012b).

[235] The other problem was that Read's account uses propositional quantification; and if this is legitimate, there are versions of the Liar paradox which finesse the machinery of the Bradwardine solution. Nothing in Read's present paper addresses this issue. I note that in (2008) Read addresses a version of the Liar paradox given by Tarski, which uses propositional quantification and the operator 'says that'. His comments there do not apply to the version I give, which does not use such an operator.

[236] As given in IC, 1.8.

In the first of these, $D(x, y)$ is the denotation predicate (x *denotes* y), and t is any closed term of the language. This is the D-Schema. In the second, we assume a relabelling of bound variable to avoid any clash when the μ-term is substituted. This is the Least Number Principle (LNP): if something satisfies A then the least thing that satisfies A is such a thing. We need make no assumption about how the μ-term behaves if nothing satisfies A (though, in his account, Read in fact does).

Since the D-Scheme is the analogue of the T-Schema, I had assumed in RBL that it was this which should be Bradwardinised for a corresponding approach to the paradoxes of denotation. This would give something like:

- $D(\langle t \rangle, x) \leftrightarrow \forall y(\mathbf{Sig}(\langle t \rangle, y) \to t = y)$

though what **Sig** means in this context is somewhat unclear; and even given this, the modification does not appear to do what is required.

In the present paper, Read endorses the unmodified D-Schema (though, interestingly, he deduces this from considerations concerning signification). What then is supposed to break the paradox is the Bradwardinised LNP. This now becomes:[237]

- $\exists x \forall P(\mathbf{Sig}(\langle \mu y A \rangle, P) \to Px) \to A_y(\mu y A)$

Here, the upper case variables range over properties, and $\mathbf{Sig}(\langle t \rangle, P)$ means that the term t signifies the property P. There is already a notable movement from the Bradwardinian theory here, since we are no longer modifying semantic principles, but the behaviour of a much more general piece of logical machinery: descriptions. Moreover, there is another significant departure. In the Bradwardinian theory of truth the entities signified by a sentence are of the same syntactic category. The same is true of a Bradwardinian account of the Heterological Paradox, where open sentences signify properties.[238] In the present case, terms signify properties—something of a different syntactic category.

Anyway, given this machinery, if '$\mu y A$' signifies (were to signify) just one property, which it possesses—any property whatever; it need have nothing to do with the property A intuitively specifies—then one can infer that

[237] I have changed Read's notation to bring it in line with the conventions used in the present essay.

[238] See both Read's paper and mine.

$A_y(\mu y A)$. However, if it may signify other things, we cannot; so the paradoxical argument is blocked.

However, why should we suppose that terms have multiple significations? Indeed, why should we suppose that they have a signification at all? Read points out that a descriptive term has a sense. Thus 'the most ignorant and stupid president the United States has ever had', has a sense which picks out its bearer. (No prizes.) We may therefore think of the signification of a descriptive term as its sense, that is, in effect, the property specified by the open sentence used in its construction. But signification is supposed to be a feature of all terms, not just descriptive ones. What does 'Aristotle' signify? Assuming that 'Aristotle' is a rigid designator in the sense of Kripke, it has no sense.[239] Read tells us that this includes at least the property $\lambda x x = \text{Aristotle}$. Quite generally, Read tells us, any term t signifies the property $\lambda x\, x = t$.

But given that this is part of the signification, why do we need to suppose that the signification of a descriptive term is something else *as well*, given by its sense? Read postulates that if a term signifies some property, X, it signifies any property, Y, such that anything that is X is Y (Clo). If this were the case, the term t would signify the property $\lambda x(x = t \vee x = \text{Donald Trump})$, which is certainly different. Clo is the analogue of what Bradwardine assumes about whole sentences (that if B follows from A, and a sentence signifies A, it signifies B). But even given that closure under signification is the case for closed sentences—and why should one suppose that? As far as I understand it, Bradwardine just postulates this—why suppose that it holds for other grammatical categories?

And the answer had better not be that otherwise Berry's paradox would give us a contradiction. One can turn *any* principle involved in the generation of a paradox into a conditional with some antecedent condition, and then use the paradox argument to infer that the condition is not satisfied. This is cheap. What we need for a genuine solution is an independent argument that the condition is not satisfied—in this case, that a term has multiple significations.

But there are other problems with the solution. Grant that this approach blocks the Berry argument. It also blocks every other argument in which we use the least number operator—and all other kinds of descriptions.[240] That's

[239]Kripke's arguments in *Naming and Necessity* to the effect that proper names are not descriptions of any kind are well known, and need no rehearsal here.

[240]Read notes that the strategy blocks many other paradoxes: König's paradox, Berkeley's paradox, Hilbert and Bernays' paradox. In a sense, this is not, therefore, surprising.

too much: we reason correctly using descriptions all the time. Given Read's approach, to deploy any description, $\mu x A$, we need to establish that the modified description principle can be applied. How do we do this? Proving that $\exists x A$ will not do, as Berry's paradox shows us. What we have to prove instead is that $\exists x \forall P(\mathbf{Sig}(\langle \mu y A\rangle, P) \rightarrow Px)$. The fact that this is not the case when the properties that '$\mu x A$' signifies are not mutually consistent might suggest that we can take it to hold if they *are*. But proving consistency is, we know, hard, and in general highly non-effective. Moreover, whatever it takes to show that this condition holds, we cannot even make a start on it till we know what properties the term $\mu x A$ signifies. Nothing in the story so far tells us this.[241] Without a resolution of these issues, the job is at best half done.[242] In truth, this was already an issue with the Bradwardine solution to the liar paradox. We frequently use the right-to-left direction of the T-Scheme. (Everything true is found in the *Bible*. We should take an eye for an eye and a tooth for a tooth. So this is found in the *Bible*.) But, in the present context, the matter is much more acute, simply because we are dealing with a piece of logical machinery that is topic-neutral.

Finally, and in any case, the Bradwardinian machinery would, in the end, seem to be beside the point. Grant that the least number operator works in the way that Read says that it does. There appears to be a perfectly intelligible neighbouring operator which delivers paradox. Let me illustrate using the Berry case. The simple combinatorial argument in Berry's argument assures us that, for a certain condition, B, with one free variable x (x is number not definable in such and such a way), $\exists x B$. By the properties of natural numbers, there is a least such. Now, never mind what the term $\mu x B$ signifies, we know that there is a unique thing satisfying B. Fix on this number, and give it a name. If you like, call this $\mu^* x B$. Then, if B specifies a purely extensional context (as it does in the Berry case)—that is, one where the only thing relevant to whether an object satisfies it, is the object itself, nothing to do with the way that it is specified—then by the very construction, $B_x(\mu^* x Bx)$. We may simply run Berry's argument for this.

[241] Read tell us that $\mu x A$ signifies $\lambda x A$, and anything that this property entails. But that cannot be all. $\mu x A$ signifies $\lambda x A$ and $\lambda x\, x = \mu x A$, and these do not entail each other if we are not entitled to assume that $A_x(\mu x A)$.

[242] A dialetheic approach to the paradoxes of course faces a similar issue. This is the problem of "classical recapture". Thus, applications of the disjunctive syllogism are not valid; yet we frequently use the syllogism unproblematically. How so? There is a substantial literature on this (starting with IC, ch. 8, and IC2, ch. 16).

Let me now turn to another matter raised by Read: the Principle of Uniform Solution (PUS) and the Inclosure Schema. Berry's Paradox is a touchstone for proffered solutions to the semantic paradoxes, such as Read's.[243] The Principle of Uniform Solution (PUS) says: same kind of paradox, same kind of solution (BLoT, 11.5, 11.6). So if two paradoxes which are of the same kind are given different solutions, these cannot (both) be correct.

Since the Liar Paradox and Berry's Paradox would clearly seem to be of the same kind, it should be the case that solutions such as Read gives are of the same kind. Whether Read's solution to the two paradoxes satisfies this criterion is somewhat moot, as I have noted. Certainly, the machinery deployed in both cases is the same. However, the natural analogue of Read's solution to the Liar paradox qualifies the D-Schema in the way that the T-Schema is modified; and this, his approach does not do. However, what is at issue here is what counts as the same kind of solution, and this is a somewhat murky depth we need not plumb here.[244]

Read is, in any case, not persuaded by the PUS. He claims (§2) that the converse principle (same kind of solution, same kind of paradox) is 'much more plausible'. Moreover, one can contrapose the PUS: different kind of solution, different kind of paradox. Whilst (perhaps) logically equivalent, this suggests applying the principle differently. We take solutions to provide a criterion for individuating paradox kinds. Thus, the mere fact that two paradoxes have different kinds of solution shows, *ipso facto*, that they are of different kinds, 'the possibility of a separate solution bringing out their different character'.

Now, first, I take the converse of the PUS to have very little plausibility. The fact that two paradoxes have the same kind of solution most certainly does not show that they are of the same kind. Thus, a dialetheist may hold that the liar paradox and paradoxes about the instant of change have a dialetheic solution.[245] This hardly shows that they are the same kind of

[243] But also Field's. See §14.5 above.

[244] Read notes (§6) that my solution to the Hilbert and Bernays paradox involves faulting some traditional laws of identity, and so is of a kind different from my solution to the other paradoxes of denotation. This is not quite right. The solution is not to amend the laws of identity: that is just a consequence. The dialetheic solution to the usual paradoxes of self-reference is to suppose that a sentence may have more than one truth value (that is, Fregean reference); the solution I give to the Hilbert and Bernays paradox is to suppose, analogously, that a term may have more than one referent. Again, this takes us into the question of what counts as the same kind of solution.

[245] Paradoxes of the instant of change were well known to the medievals under the title

paradox: self-reference has nothing to do with the instant of change.

Second, the methodology of individuating paradoxes in terms of their solutions is seriously flawed. Thus, consider the liar paradox. The self-reference required for this can be obtained in many ways: with a demonstrative (*this*); with a definite description (*the first sentence...*), gödel coding, and so on. Suppose that one were to give a solution to the version where demonstratives are used, deploying a theory according to which a demonstrative that is self-referential is ungrammatical; but that one were to give a solution to the version where a description is used, deploying a theory according to which the T-Schema fails for sentences containing descriptions. This would be bizarre. We have a sense that it is the same thing that is going on in these different formulations of the liar paradox, however self-reference is achieved; and focussing on the mode of self-reference is missing the nerve of the paradox.

Of course, how to articulate this sense of 'the same thing' is no easy matter. That is exactly the point of the Inclosure Schema.[246] This is a schema involving an operator ("the diagonaliser") which, when applied to a bunch of objects of a certain kind delivers a novel object of the same kind. And when one sees how it does this, one understands how it is that a contradiction will be generated at the limit, when the operator is applied to the totality of all such objects. The mechanism which produces the contradiction is, as it were, revealed, and one understands *why* the paradox arises. Moreover it is not just I who think that the Inclosure Schema does this. Recall that the schema is just a tweak of one proposed by Russell (1908), where he explains that it exposes the mechanism which generates the paradoxes of self-reference. Of course, the diagnosis of why this kind of paradox arises does not determine a solution. That is another matter entirely. For Russell, it was an enforcement of the Vicious Circle Principle; for a dialetheist it is accepting the contradiction at the limit. But as the PUS says, the same kind of solution is required, whatever that is.

Finally, in §6 Read raises the interesting question—which, oddly, had not occurred to me before—of whether Hilbert and Bernays' paradox fits the Inclosure Schema. At present I have found no way to show that it does so. The problem is that the things that fit the Inclosure Schema are limit paradoxes. We have a totality, and an operation on subsets of that totality, which

of 'incipit and desinit'. On a dialetheic solution to these, see IC, ch. 11; and on medieval connections, see Priest (201+g).

[246] BLoT, 11.5, 11.6 and BLoT2 17.2.

gives rise to a contradiction when things are pushed as far as (im)possible. Hilbert and Bernays' paradox doesn't seem to be a limit phenomenon of this kind. The contradiction involving the fixed point just doesn't appear to be something that happens at the limit of some totality. Until now, the major example of a paradox of self-reference which does not fit the Schema is Curry's paradox. Whether it is in the same family as the other standard paradoxes of self-reference is a vexed question.[247] But at any rate, Curry's paradox does not seem to be a limit paradox either. So maybe now we have two examples of this kind.

20 Restall: A Tale of Two Negations

Greg Restall and I have also been friends and colleagues for many years. Over these years, I have probably discussed logic with him more than with anybody else. And though his views on logic differ from mine in many ways, I have learned much from our discussions, and his technical and philosophical insights.

Restall's paper takes us into the world of negation. The system of First Degree Entailment (FDE) is, as he says (§2), the most natural, straightforward, and elegant way of handling truth value gaps and gluts. It also provides a stable basis for extensions to theories of modality, conditionality, and other topics. The system is now some fifty years old, and is very well understood. As Restall's paper shows, however, there are still novel things to be learned about it.

As he notes, there are two well known semantics for FDE.[248] According to one—the relational semantics—there are two truth values (*true* and *false*), and sentences may have two, one, or none, of these.[249] The other semantics is a modal semantics, and uses the Routley $*$ operator.[250] I have always pre-

[247]On which, see Priest (2017), §15.

[248]See INCL, ch. 8.

[249]Equivalently, one may formulate this as a four-valued logic, with values *true* (only), *false* (only), *both*, or *neither*.

[250]These have come to be known as the 'American plan', and the 'Australian plan', respectively. As far as I know, the terms first appeared in print in Routley (1984). I have never cared particularly for the terminology. It is true that it bespeaks *something* of the origins of the ideas. But it seems to me as inappropriate as calling Frege's and Russell's ideas, the German plan and the English plan for classical logic.

ferred to first approach, because of its transparent conceptual simplicity.[251] Whatever its technical versatility, the philosophical meaning of the Routley $*$ has, however, always been somewhat opaque, as is the matter of why a world-shift should poke its nose into the truth conditions of negation.[252]

As Restall shows, both semantics can be augmented to produce a second negation satisfying the same inferential rules as the usual FDE negation. The two negations interact in quite different ways in the two semantics, though. The matters concerning how to interpret the machinery which I noted in the last paragraph are on display in the constructions.

Truth and falsity are pre-theoretical notions; one might call them 'folk notions'—though doubtlessly the folk don't pay too much attention to their details. The thought that truth and falsity are exclusive and exhaustive is natural enough; but so are the thoughts that they might not be. It surfaces in many metaphysicians. Thus, Aristotle argued in *De Interpretatione*, ch. 9, that contingent statements about the future are neither true nor false; Hegel argued in his *Logic* that motion—amongst many other things—realises contradictions; then there are the many theorists of vagueness who take the borderline zones of vague predicates to deliver truth value gaps or gluts.[253] It also surfaces in the thought of the folk themselves, as XPhi studies make clear.[254] It is exactly, these possibilities concerning truth and falsity that the relational semantics make manifest.[255]

In Restall's extension of the relational semantics for FDE (§3), we still have "good old fashioned" falsity; but he adds *another* notion of falsity, which behaves in a parallel fashion.[256] One might naturally think (as I did): what

[251]See IC, chs. 5, 6, and IC2, 19.8. True, I do not accept the *neither* possibility philosophically; but that does not bear on the present issue.

[252]Restall himself goes some way towards addressing these matters in his (1999).

[253]On these, see, e.g., INCL2, 7.9, 11a.7, Priest (1990), Priest (2010b), and Fine (1975), respectively.

[254]See, e.g., Ripley (2001) and Alxatib and Pelletier (2011).

[255]Restall observes (§1) that though both truth and falsity are involved in the semantic conditions of the connectives, validity is usually defined in terms of only truth preservation forwards. It might be natural to add a clause requiring falsity preservation backwards. In fact, for FDE, this would not change the validity relation at all (IC, 8.10, ex. 8). It would for LP, since we would then no longer have $\vDash A \vee \neg A$, though we would have $B \wedge \neg B \vDash A \vee \neg A$. However, if one requires a consequence relation for which falsity preservation backwards is important, one can always define one, $\vDash^{\#}$, in the obvious way: $A \vDash^{\#} B$ iff $A \vDash B$ and $\neg B \vDash \neg A$ (with its natural generalisation to the multiple premise and conclusion case).

[256]Though there are other possibilities for adding a second notion of falsity, as he notes.

on earth is that supposed to be? The philosophical meaning of this second notion is opaque. One might, of course, raise a skeptical question. How does one know that it is the first falsity that represents the usual notion, and not the new notion? But that point can be set aside. For each negation, on its own, behaves in exactly the same way. In a sense, as Restall notes, they collapse into each other. Differences emerge only with their interaction. We are still faced with the question of why one should have have *two* structurally parallel notions of falsity at all.

Compare this with Restall's extension of the $*$ semantics (§4)—which is clearly the more interesting extension mathematically. One is not at all inclined to ask what the second $*$ operator means, precisely because it was not clear what the original one was supposed to mean in the first place. There seems to be no particularly good reason why the machinery of stars should not be multiplied *ad lib*.

None of this bears on Restall's interesting technical investigations and results, of course. I am just using his construction to bring out philosophical issues behind the two technologies for negation—a matter that Restall does not broach in his paper.

21 Shapiro: *L'Affaire Gödel*

21.1 Background

Stewart Shapiro and I have been friends for many years. For over a decade we were both Arché Professorial Fellows at the University of St Andrews, and we had many fruitful and enjoyable discussions over these years.[257] His paper here revisits *l'Affaire Gödel.*[258] I think it will help the discussion of this to put it into the context of the history of the whole matter.

In 1971 I attended the Bertrand Russell Memorial Logic Conference in Uldum.[259] During this, there was a talk on Gödel's Incompleteness Theorems by Moshé Machover. A central point of his discussion was how it could be

[257]He is fond of telling the following story. In our visits to St Andrews we frequently shared an apartment. Often we would start a discussion at supper, and take it up again at breakfast. He is an evening person, and I am a morning person. He would get the better of the discussion in the evening, and I would get the better of it the following morning.

[258]For more on the matter see the comments on Field, §14.4 above.

[259]Actually, this was the first conference I ever attended. The results of the conference were subsequently published as Bell, *et al* (1973).

possible that there are things which are unprovable, but which we could yet know to be true. The thing in question was, of course, the sentence that says of itself that it is not provable, that is, a sentence, G, of the form $\neg Prov\langle G\rangle$, where $Prov$ is the proof predicate for the axiomatic arithmetic in question.[260] For the rest of these comments I will refer to this—perhaps somewhat inappropriately, given the context—as the 'undecidable sentence'. This sentence changes from theory to theory, of course. The theory in question will, I hope, be clear from the context in what follows.

Anyway, Machover's problem piqued my curiosity. The problem was not, of course, that there are things that can not be proved in some system, but proved in another. The point was that we can recognise the truth of the undecidable sentence for, say, PA by means which are in some sense implicit in what we can already prove. One way one might make the point is this. PA uses an axiom schema of induction. But whatever intuition supports the schema supports, equally, the second-order version of induction, where one merely replaces the schematic variables with a second-order variable. Yet in second-order arithmetic one can prove the first-order undecidable sentence. Similarly, Dummett suggests that reasoning by mathematical induction is constitutive of our concept of natural number, but:[261]

> once a system has been formulated, we can, by reference to it, define new properties, not expressible in it, such as a true statement of the system: hence, by applying induction to such new properties, we can arrive at a conclusions not provable in it.

What is going on here?

As I thought more about the phenomenon it seemed to me that the essence of the matter was how we show that the undecidable sentence is true in the standard model of arithmetic. (As is evident to anyone who has ever thought about soundness proofs, showing that the axioms of PA hold in the standard model invokes the very claims that are stated in the axioms.) Since we are dealing with the standard model, we might just as well talk about truth *simpliciter*. Hence, to carry out such reasoning, we need a language with a truth predicate. And of course, if the undecidable sentence is indeed provable, the theory is inconsistent. So we have a contradiction on our hands.

[260] Standard proofs of the fixed point theorem deliver only a G equivalent to $\neg Prov\langle G\rangle$. But if the function symbols available include one for diagonalisation, we can obtain a literal identity. (See IC 3.5.)

[261] Dummett (1978a), p. 195. Dummett's view is discussed in IC, 3.2.

Paraconsistency was therefore required. Trying to tie down this thought was what lead to the discussion of Gödel's theorem in §2 of the 'Logic of Paradox' (Priest (1979)) and two later places.[262] In what follows, let me call the formulation of the argument in these places, the *original formulation.*

So, *pace* the introduction to Shapiro's paper, the aim was never to have a 'complete, decidable, yet sufficiently rich mathematical theory'. I did (and do) not find a problem with the thought that there are true mathematical claims that cannot be proved. What worried me was the thought that there was a particular unprovable claim that we could know to be true. And there was never a suggestion that arithmetic should be decidable—welcome as it might be if this were the case.[263]

At any rate, I now think the the original formulation overshot the mark. For there to be an undecidable sentence in the first place, the theory in question had to be axiomatisable (or at least, arithmetic). I argued that naive mathematical proof (all of it) was, in principle, axiomatic. This was unnecessary. All that needed to be axiomatisable was the fragment of it in which the argument for the truth of the undecidable sentence was carried out—a much more modest claim.

Anyway, and to return to the history: The original formulation of the argument made no attempt to spell out the details of what an inconsistent arithmetic which can prove its own undecidable sentence would be like.[264] Things changed when I became aware of the potential for applying the work of Meyer on relevant arithmetic,[265] and especially once the techniques of the Collapsing Lemma because available in Priest (1991). The resulting

[262]I revisited the matter in Priest (1984), §§5-7. IC, ch. 3, is a slightly topped up distillation of these two discussions.

[263]I note that the notion of completeness is ambiguous in this context. Completeness might mean that, for every A, either A or $\neg A$ is provable; or it might mean that everything true is provable. The inconsistent arithmetics constructed in INCL2, ch. 17, are complete in the first sense; they may not be complete in the second; that depends on what actually *is* true. The axiomatic theories constructed there are finite, and so decidable. (Later, Paris and Sirokofskich (2018) showed that there were infinite decidable models.) But decidability was simply a route to showing axiomatisability. I presume that there are complete (in the first sense) axiomatisable but non-decidable inconsistent arithmetics, though I know no proof of this.

[264]That was one of the tasks that Shapiro (2002) determined to undertake. It was a good shot, but turned out to have problems, as Shapiro notes in §§3, 4 of the present paper.

[265]As spelled out in Meyer and Mortensen (1984). (The first edition of IC was essentially finished in 1983. It did not appear until 1987 because of the difficulty of finding a publisher, as the preface to the second edition of IC explains.)

implications were spelled out in Priest (1994a) and (2003).[266] These formed the basis of the material in IC2, ch. 17.

Two things are notable about these inconsistent arithmetics. First, the working assumption about inconsistent arithmetics in the original formulation was that the underlying logic of the theory was a logic with a detachable conditional. The inconsistent models of arithmetic showed how the result—reasoning to the truth of the undecidable sentence—could be achieved when the underlying logic was LP, and so doesn't need a detachable conditional.[267] This means, of course, that reasoning within the theory, one does not use *modus ponens*. Given whatever axioms there are, the rules of LP deduction suffice. Since the discovery of these inconsistent arithmetics, I have tended to the view that the phenomenon at issue is best understood in the language of LP. This is, after all, the standard classical assumption as well—and if it is correct, issues about a detachable conditional fall by the wayside.

The second notable thing is that the language of the theories constructed does not, in a certain sense, contain a truth predicate. In one sense, $Prov$ is a truth predicate for the theory. For any A, either A is in the theory or it is not. Since $Prov$ represents proof in the theory, in the first case $Prov\langle A\rangle$ is in the theory; in the second case, both $\neg A$ and $\neg Prov\langle A\rangle$ are in the theory. In both cases, $A \equiv Prov\langle A\rangle$ is in the theory. But of course, this does not give us bi-deducibility for the T-Schema. However, there is a well-known construction that does deliver such a truth predicate. One can take an inconsistent model of arithmetic, add a truth predicate to the language, and conservatively extend the theory to one which verifies the T-Schema in a bi-deducibile form (as IC2, 17.3 notes). If one starts with an axiomatic theory of arithmetic and adds the T-Schema, one obtains an axiomatic theory. So the extended theory has an arithmetic proof predicate, $Prov$. One can show that $Prov\langle A\rangle \supset A$ is true in the theory (as in IC2, 17.4), as, then, is $Prov\langle A\rangle \supset T\langle A\rangle$. It does not follow that one can prove that $\forall x(Prov\, x \supset Tx)$, however. If the numbers in the model are all standard (and $Prov$ is defined in such a way that if n is not the code of a formula then n satisfies $\neg Prov\, x$) then this is true in the extended model. However,

[266] I note that the first of these deploys what is, in effect, the Hilbert and Bernays paradox about denotation. I no longer accept that argument. See TNB, ch. 8.

[267] Interestingly, an axiomatization for some of these inconsistent arithmetics using the logic A_3, which conservatively extends LP with a detachable conditional, was given by Teddar (2015).

it does not follow that it is provable.

Given these things, I think that the original formulation of matters overshot the mark in another way. First, the proof of the undecidable sentence does not require the quantified form of soundness: $\forall x(Prov\, x \supset Tx)$. Since a particular sentence is at issue, the schema $Prov\langle A\rangle \supset T\langle A\rangle$ will do just as well. Indeed, even the truth predicate itself is not necessary. The plain $Prov\langle A\rangle \supset A$ will do. Secondly, and perhaps more importantly, it is not necessary that soundness—in any of these forms—is provable from other things.[268] Even if it is a simple axiom schema, this will suffice for the proof of the undecidable sentence. The question of proving its truth from other things also, therefore, falls by the wayside.

21.2 Church's Thesis and its Application

With this background, let us now turn to the contents of Shapiro's paper. §§2, 3, and 4 provide a discussion of infelicities in Shapiro's earlier proposal of an appropriate formal arithmetic. There is little for me to comment on here.

§1 provides a number of arguments against the appeal to Church's Thesis in the original formulation of the argument (though some of them raise matters of more general import). The relevant question here is the extent to which the considerations he raises apply to a version of the argument to dialetheism which does not overshoot the mark, as the original formulation did. Shapiro raises five concerns. Let us take them in order.

Point 1: This is an objection to mathematical foundationalism—and I clearly did appeal to this in the original formulation. (This was, after all, the 1970s, and foundationalism was still the dominant view.) Mathematical foundationalism is, in fact, a view I have subsequently come to reject.[269] The foundationalism of the original formulation was used to justify the thought

[268]Happily. Since, as Field later stressed (see §14.4 above), and as Shapiro notes in §4 of his paper, natural as the argument it is, Curry paradox considerations rule this out—at least if the logic of the theory uses *modus ponens*.

[269]Indeed, Shapiro's views on the philosophy of mathematics, which he was developing in St Andrews, and which were subsequently to appear in Shapiro (2014) helped to push me in that direction, though the move was well under way before that. (See Priest and Thomason (2007).) Indeed, rereading what I wrote many years ago, I was surprised to find that the move was even prefigured in my early papers on the undecidable sentence. (Priest (1979), §IV.11, and Priest (1984), p. 171.)

that naive mathematical proof is axiomatic. This, in turn served two guarantee two things necessary for the argument:

[E] That there *is* an undecidable sentence.

[S] That what is proved is true.

Throw in the claim that the undecidable sentence is provable, and one has the sought result.

Now, granted that foundationalism in general is wrong, and that there is no reason to suppose that mathematics, *in toto*, is axiomatic, what of the fragment of it involved in the argument for the undecidable sentence? This fragment contains little more than the basic facts of arithmetic (e.g., those available in PA), plus the statement of soundness. Let us call this fragment of naive proof the *modest fragment*. The modest fragment would certainly seem to be axiomatic. That is, [E].

What about [S]? It is hard to cast doubt on the basic facts of arithmetic. So the main suspect for doubt is soundness itself. However, if we understand *proved* simply as *established as true*, this just seems analytic. True, soundness is perhaps not something that mathematicians would normally concern themselves with explicitly, but no mathematician is going to contest the claim that things that have been proved are true. (Recall that most mathematical proof is not proof in a formal system. Formal regimentation comes later—if at all.) Moreover, even if soundness is false for the system in question, we can, presumably, revise it to make it so—and the basic facts of arithmetic are unlikely to be junked in the process.[270]

Point 2: This concerns [E]. Shapiro points out that for the notion of recursiveness to make sense, we need a fixed language, and the language of mathematics is both somewhat indeterminate and changes. Again I agree; but the point seems to have no force against the modest fragment of naive proof. For the language of Peano Arithmetic will do, perhaps augmented by a truth predicate. And considerations of temporal change are irrelevant: we are dealing with how we reason here and now.

Point 3: This concerns the notion of consensus (or agreement). Shapiro points out that in many areas of mathematics, consensus is sometimes hard to achieve, and is based on a 'relatively small and finite sample of mathematicians'. First, since I have not mentioned the topic till now, why is consensus

[270]See IC, p. 46.

relevant? It is invoked twice in IC, ch. 3.[271] In both cases consensus is used to support the idea that naive proof is axiomatic. This is no longer an issue, it seems to me, for the modest fragment. (See Point 1.) And in any case, it seems reasonable to suppose that there would be consensus concerning that particular fragment.

Point 4: This also concerns consensus. Shapiro says that where there is consensus, it is based on the assumption of classical (or at least intuitionist) logic—which I am obviously in no position to endorse. Now, for a start, mathematical proof is carried out informally. (No one argues *à la Principia Mathematica*.) I doubt that most mathematicians who work outside of the foundations of mathematics (which is most of them) know (or care!) much about formal logic—classical or otherwise. They just reason in a way that seems right to them. Moreover, as far as the modest fragment goes, this concerns only the inferences of *LP* (and maybe some inferences concerning a detachable conditional and truth). There is nothing very contentious about any of these for most mathematicians.

Point 5: Perhaps the preceding observations serve to circumvent many of the disagreements between Shapiro and myself; but this may not be the case with respect to the last point, which concerns the application of Church's Thesis itself. In my original formulation of matters, this was invoked to support [E]. Shapiro objects that appeal to the thesis is not appropriate since to do so 'one must specify an algorithm, a step by step procedure for computing a value, for deciding a question, a procedure that invokes no creativity or use of intuition' (§1). Now, Church's Thesis can be put in different ways; but one standard version of it is to the effect that if a function is effectively computable, then it is recursive in the technical sense. Indeed, this is Church's own formulation of the Thesis:[272]

> We now define the notion, already discussed, of an *effectively calculable* function of positive integers by identifying it with the notion of recursive function of positive integers.

The previous discussion in question consists simply of examples. The function at issue in the present case is that which maps (the code of) a sequence of sentences to 1 or 0, depending on whether it is sound argument. And, as Church himself says elsewhere:[273]

[271]Pp. 40 (fn 3), 41.
[272]Church(1936), p. 356.
[273]Church (1953), p. 53.

> ...consider the situation which arises if the notion of proof is non-effective. There is then no certain means by which, when a sequence of formulas has been put forward as a proof, the auditor may determine whether it is in fact a proof. Therefore he may fairly demand a proof, in any given case, that the sequence of formulas put forward is a proof; and until the supplementary proof is provided, he may refuse to be convinced that the alleged theorem is proved. This supplementary proof ought to be regarded, it seems, as part of the whole proof of the theorem...

Indeed, that proof is effectively recognisable and truth is not is a central way in which the two differ, as Dummett, for example has argued.[274] Now, it may well be that this appeal to effective recognisability breaks down where consensus breaks down, but I have already dealt with this matter. (Points 3 and 4.) So I still see no problem with invoking Church's Thesis in the present context.

In any case, and as Shapiro notes at the beginning of §2, arguments for [E] are finessed if one can actually produce such a system. This was one of the purposes of the system Shapiro unsuccessfully produced in (2002). But the inconsistent arithmetics of IC2, ch. 17 do do this—though perhaps they do not provide entirely what is required, since there are many axiomatic inconsistent arithmetics, so this does not determine the system uniquely. Different considerations are required to do so—if, indeed, there is a uniquely correct such system: it could that there is an indeterminacy in the matter. Equations may go inconsistent for suitably large (*very* large) numbers; but where, exactly, we may not know.[275]

21.3 Curried Undecidability

This brings us, finally, to the interesting matter of the Curried undecidable sentence, which Shapiro discusses in §5. Shapiro claims that it shows that even in an axiomatic framework which can prove its own Gödel undecidable sentence, there are still true but unprovable sentences. That, *per se*, would not be a problem: as I have already said, the argument was never aimed at showing that all truths are provable. The problem was with a sentence we

[274] Dummett (1975b)

[275] The matter is discussed in Priest (1994b).

could know to be true, but which was not provable. How do things stand with the Curried undecidable sentence?

We assume that we have an inconsistent axiomatic arithmetic. Then there is a sentence, C, of the form $Prov\langle C\rangle \Rightarrow \bot$. The first question is how one is to understand $\Rightarrow$. If the arithmetic is one where the only conditional available is $\supset$, then C is $Prov\langle C\rangle \supset \bot$. This is logically equivalent to $\neg Prov\langle C\rangle$. So this is simply the Gödel undecidable sentence, and there is nothing new here.

Suppose, however, that we are contemplating an arithmetic that has a detachable conditional, $\rightarrow$, and that this $\Rightarrow$. If C is provable, then so are $Prov\langle C\rangle \rightarrow \bot$ and (since $Prov$ represents provability) $Prov\langle C\rangle$. Hence the theory is trivial. If it is not, then C cannot be proved. So, $\neg Prov\langle C\rangle$ is true. All this is as Shapiro notes.

What of the truth of C? Given that we are not dealing with a material conditional, we cannot simply infer that $Prov\langle C\rangle \rightarrow \bot$, i.e., that C is true. But, as Shapiro notes, there is another argument. Suppose that $Prov\langle C\rangle$, then by soundness, $Prov\langle C\rangle \rightarrow \bot$, and so $\bot$. Hence, by conditional proof, $Prov\langle C\rangle \rightarrow \bot$, i.e., C. This is essentially the $Prov$ form of the Curry paradox, and the argument must fail for similar reasons. The reason given in IC, ch. 7 is that $\rightarrow$ does not satisfy Absorption (Contraction). As is well known, in a natural-deduction context, this means that in applications of the rule of conditional proof one cannot discharge more than one occurrence of an assumption, which this argument does.[276] So is C true or not? Without a theory of the conditional and its interaction with the other machinery (and maybe even with it), one cannot say.

It remains the case that we have shown $\neg Prov\langle C\rangle$ to be true. Just as for the Gödel undecidable sentence, this should therefore be provable our formal arithmetic. Whether or not it is, will also depend on the theory and, crucially, how it handles $\rightarrow$.[277]

[276] As Shapiro observes in fn 4. Note that, for the same reason, the proof of Löb's Theorem, which is a version of Curry-reasoning, breaks down.

[277] We need to be able to encode the following reasoning. Suppose that $Prov\langle C\rangle$. Then $Prov\langle Prov\langle C\rangle \rightarrow \bot\rangle$. But $Prov\langle A\rangle \rightarrow A$, so $Prov\langle C\rangle \rightarrow \bot$, and so $\bot$. Hence we have shown that $Prov\langle C\rangle \vdash \bot$. (Note that we have not proved the corresponding conditional.) Thus, $\neg\bot \vdash \neg Prov\langle C\rangle$, and so $\neg Prov\langle C\rangle$.

22 Tanaka: Taking Exception to Candrakīrti

Koji Tanaka is one of the few logicians who shares my interests in the Asian philosophical traditions, and he brings this interest to bear in his discussion of so called logical exceptionalism. I think that the best thing I can do to address his comments it to start by saying what I take logical exceptionalism to mean—or at least, what it means in as far as it applies to me.[278]

In matters of any complexity, we have to theorise. Thus, we do this in physics and biology, but also in linguistics, history, ethics, and metaphysics. In each case, there is something we wish to explain and understand (and maybe, for certain kinds of topics, make predictions). We construct theories which do this, and then accept whichever does the best job of the matter. Of course, we may change our mind as to which one does so, as new theories are discovered, or we come to understand old theories better.

Certainly, the theories must do justice to the data concerning whatever it is we are trying to explain, though data is just as liable to be fallible as theory. But adequacy to the data alone is not enough. There may be theories that do equal justice to the data—or more likely, no (extant) theory may account for all the data. So other criteria, such as simplicity, unifying ability, extent of *ad hoc* auxiliary assumptions, etc., come into play. The theory it is rational to accept is the one which does best overall—in some way of cashing out this idea.

The general picture is familiar enough. Anti-exceptionalism about logic is the view that logic, in one sense of that word, is no exception to the picture.[279] Logic, in this sense, is a theory about what follows from what, and why. It should be stressed that this is not, in any simple sense, a descriptive theory about how people *actually* reason. That is a matter for cognitive psychology. It is about the norms that govern correct reasoning.[280]

Logicians have been constructing such theories—which disagree on many

[278] As explained and defended in Priest (2014b), (2016a).

[279] Just to be clear, this doesn't have anything much to do with Quine's 'change of logic, change of subject' argument, as Tanaka suggests (§1). That's another matter. See DTBL, ch. 10, esp. 10.9.

[280] In one sense of that word. Thus, Harman (1978) prefers to use the word 'reasoning' for what would now be called belief-revision. Naturally, there are connections between these two senses. When juggling one's beliefs, if one is aware that B follows from A, then one shouldn't accept A but not B. Exactly how one cashes out this idea is no easy matter, however, as MacFarlane (2004) shows.

matters—as long as there have been logicians. The theories are contestable, fallible, and modern accounts, at least, apply quite sophisticated mathematical ideas. Seen in this way, the epistemology of logic that emerges is quite distinct from a traditional *a priori* account of the Kantian kind. In this, the laws of logic are available simply to reflection, are certain, and not rationally contestable

An obvious question that arises in this context is what kind of data it is which is relevant in logical theorising. I take these to be those simple inferences (and perhaps inference schemas) that strike us as correct; and I suppose that these are *a priori* in one sense of that term. One does not have to go and look or listen to determine their apparent acceptability; one just has to think.[281] But, as ever in theorising, what appears to be so may not be so.

With this background, let me now comment on Tanaka's paper. First, he asks at the end of §3 what, exactly it is that makes my view different from a traditional *a priori* view of logic. I hope that the answer to that question is now clear.

Next, Tanaka suggests that my account is, or can best be seen as, a form of the *lokaprasiddha* (common sense) account of Madhyamaka Buddhism. Now, how, exactly, to understand this view is somewhat contentious, but this is not the place to enter into scholarly disputes about this matter; so let us just accept Tanaka's interpretation. According to this, truth is simply what people accept. ('Truth is nothing more than what people on the street assent to and knowledge is nothing more than what they think' §4.)[282] In particular, a claim about validity is true if people accept that it to be so. Now, it is certainly the case that data against which we judge our logical theories is determined by the inferences that strike us as valid. But that is about where what is in common between this view and mine ends. In particular, it is not the case that 'whether or not an inference is valid is ... just a matter of what people in the street would accept' (§4). The results of logical theorising may well, and frequently do, overturn the views of "people in the street". Nor is it true that 'we may not have any sustained reason for

[281] I suppose that in some sense this is an empirical survey—albeit one with a single subject! I don't need to find out what others think, any more than I, as a native speaker of English, have to consult others to determine whether 'the cat sat on the mat' is grammatical.

[282] For a rather different interpretation of truth in Madhyamaka, see Priest, Siderits, and Tillemans (2010).

why we believe [a given inference] to be valid' (§4). The theories logicians develop—be they model-theoretic, proof-theoretic, or whatever—provide just such reasons. And it would just be a travesty of my account to say of it that (§4):

> there is no need to analyse logical concepts, the notion of validity or anything. All there is to logic is what can be expressed by things like 'This inference looks good to me', or 'that inference strikes me as valid'.

Even if one must start from "common sense" the aim of theorising is always, as it were, to get behind appearances. This is so in all forms of theorising, in logic as elsewhere.

Indeed, the *lokaprasiddha* account is in dire danger of collapsing into relativism about truth, since it 'reduces truth and knowledge to mere opinions and beliefs' (§5). I would reject such a relativism entirely. I take it that there are objective truths about validity, just as much as there are about physics. Indeed the whole point about theorising is to delve into the matter of what *is* true in this sense, and to deliver us our best (current) view about what this is.

Let me, finally, turn to the connection between these methodological issues and paraconsistency. Tanaka says that my methodological views on logic are 'a part of [my] argument for paraconsistent logic' (§2). In a sense this is true. The sense is that they mount a case against those historically benighted people who think that so called classical logic, *qua* theory—or any other theory, for that matter—is god-given and uncontestable. It is not true that I think that applying this method *must* deliver the conclusion that a paraconsistent logic is (currently) the most rational theory to accept (though as a matter of fact, I think it does). What comes out of applying this method can be determined only by, well, applying the method.[283] One has to look at how well each of the relevant theories performs on each of the theoretical desiderata, and on their aggregation.

A couple of issues that Tanaka mentions here are relevant. First, I think it quite true, as he says, that most people who have never studied any logic find bizarre claims to the effect that instances of Explosion are valid. This is surely a black mark against classical logic; but it is one that it does not

[283] A case-study in applying this can be found in Priest (201+h). No one has yet undertaken the daunting task of applying the method to logic quite generally.

have much difficulty digesting. The validity of the inference follows from a theory that is simple and powerful. The inference turns out to be vacuously valid; and the intuition that it is invalid can simply be explained by saying that such a kind of validity does not normally occur to people.

More telling are examples where the premises are thought to be true, and the conclusion is not. If a person can be brought to agree that there are such situations then they will surely judge that Explosion is not deductively valid. After all, not everything is true. And truth-preservation appears to be a rather minimal necessary condition for deductive (as opposed to non-deductive) validity. So any inference where the premises are taken to be true, and the conclusion is taken to be untrue will strike one as (deductively) invalid.

The point of contention here will (of course) be whether there *are* situations in which a contradiction is true. The example of visual illusions that Tanaka cites are not of this kind (*contra* the view he attributes to me in §6). These are, after all, *illusions*.[284] A contradiction may *appear visually* to hold, but it does not really do so. Thus, in the waterfall illusion, the object on which one is focussing is not really both moving and not moving; and we know this to be so.

Examples which *are* of this kind are those which dialetheists standardly cite.[285] Of course, such examples may well be contentious; and at issue will be how one might explain away the apparently contradictory nature of the situations. One gets a glimpse here of the enormity of the task of evaluating logical theories. One cannot, in the last instance, disentangle such evaluation from the evaluation of theories in semantics, metaphysics, the theory of norms, and numerous other areas.[286]

23 Tennant: Through an Inferentialist Telescope

Neil Tennant takes us into the world of what he nowadays calls Core Logic, and casts an inferentialist eye on *LP*. Our interest in each other's work goes back to the 1970s, when we both developed ideas in paraconsistency and in

[284]DTBL, 3.3.
[285]For a short catalogue of these, see §11.1 above.
[286]For some further discussion, see Priest (201+i).

logics which are, in some sense, relevant. I developed *LP*; he developed his cut-free logic. Philosophical interest in cut-free logics, and more generally substructural logics, is now reasonably commonplace.[287] It was not in those days. And over the years, Tennant has articulated his view, throwing into the mixture anti-realism, inferentialism, and paradox.

Behind his paper, there is an important difference between us concerning meaning. He subscribes to an inferentialist account of this; I subscribe to a truth-conditional account. For an inferentialist, the meanings of the logical constants are determined by rules of inference. Matters can be set up equivalently in terms of either natural deduction or a Gentzen-style sequent calculus. For the sake of definiteness in what follows, let me talk in terms of the latter. Each logical operator comes with a pair of rules, one of which shows how to introduce sentences containing the operator on the left-hand side of the sequent connector, and the other of which shows how to introduce them on the right-hand side. Moreover, the rules have a certain balance (or harmony, as it is often called), which allows for an appropriate cut theorem (or, in the case of natural deduction, a normal form theorem). This account of meaning goes naturally (though not invariably) with a proof-theoretic notion of validity, and a view of truth as warranted assertibility. For someone who subscribes to a truth-conditional account, the meanings of the connectives are given by their truth conditions (generalised to truth-in-an-interpretation-conditions, or, for some logics, truth-at-a-world-in-an-interpretation-conditions). This account of meaning goes naturally (though not invariably) with a model-theoretic notion of validity, and a more robust notion of truth.

That difference is a deep philosophical one, and this is not the place to discuss it. (I have had my say on the matter in DTBL, ch. 11, and I shall not repeat it here.[288]) I note only the following. Many of Tennant's critical remarks concerning *LP* turn on the fact that its system of proof does not satisfy the sort of desiderata one would want if one is an inferentialist. Indeed it does not; but for those who prefer a model-theoretic account of validity, such constraints are of no import. The proof system is just a way

[287] For example, with the work of Cobreros *et al.* See §§9, 11 above.

[288] Though let me say that I would not go to the wall over this matter. If an inferentialist account turns out to be correct, then so be it. There are certainly proof-theoretically defined logics that are appropriate for dialetheism. Indeed, since Core Logic (or Classical Core Logic) is paraconsistent, then, depending on many other moving parts, that might even turn out to be the most appropriate one.

of characterising the notion of validity in a combinatorial fashion.[289]

With these background comments, let me turn to four more specific matters.

First, as Tennant points out, and as is well known, the material conditional in LP, $A \supset B$, that is $\neg A \vee B$,[290] does not satisfy detachment. Ever since early years of work on LP, it has been a standard thought that the language needs to be augmented by a conditional that does,[291] and so the semantics need to be extended to accommodate this. The working assumption has usually been that the conditional is one of an appropriate relevant logic. Proof-theoretically, much of the material pertaining to relevant logic has now been subsumed under the investigation of substructural logics, which contains many results concerning a proof theory of the kind that inferentialists like. However, this is not the place to go into that matter.[292]

Next, if someone thinks that some particular logic, L, is the correct one, it clearly behoves them to give the metatheory for L (in whatever form that takes) in a way that is logically kosher. In particular, if moves are made in the metatheory which are not valid in L, there had better be an appropriate story to tell about this. As Tennant notes, the standard model-theoretic semantics for LP uses things like the disjunctive syllogism,[293] which is not valid in LP. This has motivated the many discussions of "classical recapture".[294]

In fact, the situation is even more acute for someone such as myself. For a model theoretic semantics is carried out in set theory, and this would standardly be taken to be ZF; but such a set theory is not correct if one subscribes to a paraconsistent set theory. So, again, it behoves someone like me to show either how the standard model-theoretic reasoning can be accomplished in a paraconsistent set theory (using a conditional which detaches,

---

[289] Thus, the puzzles (§2.3) and potholes (§3) that Tennant perceives are ones that appear only if one is driving on the inferentialist side of the road.

[290] I note that one can simply define $\supset$ in this way. If one wishes to take it as an undefined symbol, it can be characterised by the two-way natural deduction rules (or the equivalent for a sequent calculus):

$$\frac{\neg A \vee B}{\overline{A \supset B}} \qquad \frac{A \wedge \neg B}{\overline{\neg(A \supset B)}}$$

These rules will not, of course, satisfy an inferentialist of Tennant's kind. But there is nothing wrong with them from a model-theoretic perspective of validity.

[291] A possibility that, oddly, Tennant does not mention.

[292] For an excellent exposition of the area, see Restall (2000).

[293] At least if the conditional of the metatheory is supposed to be $\supset$.

[294] Starting with Priest (1979), §IV, and IC, ch. 8.

since the material conditional is of no use in this regard), or explain how the ZF reasoning itself is acceptable. Both approaches are to be found in the literature. I prefer the second. Since I have discussed this in §3.2 above, I will not go into the matter again here.

Third, Tennant claims that a virtue of his account is that it provides a 'faithful and homologous' formalisation of informal arguments, which a suitable proof-theoretic account, such as his, does, and which LP—and more generally, a model theoretic account of validity—does not do. Now, neither of us thinks that a theory of logic should be a theory of the way that people actually reason. That is a matter for the cognitive psychology. Rather, a logic should undergird good reasoning. For contemporary logicians, the paradigm of good reasoning has always been mathematics.

Next, I note that when people reason—and reason correctly—they reason informally, even in mathematics.[295] Arguments are not laid out in some natural deduction system or sequent calculus. Such would be far too prolix. Those for whom this is important take it that the arguments *could* be regimented in an appropriate fashion. Tenant claims that the regimentation should be given in terms of the rules beloved by an inferentialist proof theorist. However, I see no reason for this unless one supposes in advance that such is the correct understanding of validity. Indeed, the original and most famous regimenters of mathematical reasoning were Frege, in *Grundgesetze*, and Russell and Whitehead, in *Principia*; and their regimentations were clearly nothing like the kind Tennant endorses.

Of course, we will often want to put a piece of reasoning under the microscope. This is highly useful in assessing the correctness of a complex piece of reasoning. However, what is at issue here is whether the steps in the argument are valid. This is the business of a theory of validity, and it does not have to be an inferentialist one. A model-theoretic one will do just as well.

Finally, though Tennant subscribes to a paraconsistent logic, he is no dialetheist. In particular, he is not a dialetheist about the semantic paradoxes.[296] Though he endorses the T-Schema in its bi-deducibility form, he argues that the failure of Cut (or normalisability) avoids the dialetheic conclusion. He points out that one can establish that $\vdash \neg Tl$ and that $\neg Tl \vdash \bot$, where l is the liar sentence; but one cannot deploy Cut to infer $\vdash \bot$. However,

[295] Tennant invokes the fact (§1) that in Priest (1979) I refer to naive proofs. As the context, I would hope, makes clear, this was the sort of proof I was talking about.

[296] For this and what follows, see Tennant (2015).

in Core Logic, the second fact delivers $\vdash \neg\neg Tl$. Even if we cannot infer $\vdash Tl$ for reasons of an intuitionist kind, it remains the case that we have established a contradiction. If the deductions are kosher, we have dialetheism. Moreover, the failure of the inference to $\perp$ is just what the dialetheist needs so prevent the spread of contradiction.[297]

24 Verdée: Making LP Relevant

Let us now turn to the paper by Peter Verdée, and the issues of conditionality which it raises.

The material conditional of LP, $\supset$, does not detach. It is therefore an obvious idea that the language needs to be augmented with one that does. Such a thought has been pursued since the inception of LP.[298] A plausible thought is that the conditional should be that of an appropriate relevant logic. If the conditional is to be the one used to formulate principles which generate the paradoxes of self-reference, such as the T-Schema and the naive comprehension schema of set theory, then not all such conditionals are appropriate. Thus, the Absorption principle, $A \to (A \to B) \vdash A \to B$, will deliver triviality, in the shape of the Curry paradox. Relevant logics which contain this or related principles, such as R and Verdée's own conditional, are not, therefore, appropriate. If, however, dialetheism is not to be deployed to solve these paradoxes; or if it is, but the conditional of the naive principles is a non-detachable conditional such as $\supset$,[299] there is no problem with a relevant logic containing such principles.

As something of an aside, let me note the following. It is sometimes objected to a dialetheic solution to the paradoxes of self-reference that it is not uniform: the solution to the Liar paradox rejects Explosion; the solution to Curry's paradox rejects Absorption. Since these paradoxes appear to be of the same kind, they should have the same kind of solution (the Principle of Uniform Solution). However, for a start, if the conditional of the Curry sentence is $\supset$, then the Curry paradox is solved by rejecting the disjunctive syllogism, $A, \neg A \vee B \vdash B$. In LP this is a very simple equivalent of Explosion. Hence the solutions to the two paradoxes are exactly the same. If,

[297]One may make a similar point about the cut-free ST system of Cobreros, *et al.* See §11.2 above.
[298]Priest (1979), §4, IC, ch. 6.
[299]As suggested by Goodship (1996), and discussed in Priest (2017).

however, the condition of the Curry sentence is a detachable conditional, $\to$, this is no longer the case. But there is no *a priori* reason why *different* sorts conditionals should require the *same* sorts of solution.[300] And there are, in fact, good reasons as to why the Liar paradox and Curry's paradox with a detachable conditional are *not* of the same kind.[301]

Anyway, and to return to Verdée's paper, he suggests a novel kind of relevant conditional. The strategy used can be applied to any consequence relation, but Verdée's main aim is to apply it to LP. The strategy comes in two stages. The first is to take a consequence relation and use it to define a sub-relation which is relevant (presumably in usual variable-sharing terms, though Verdée never specifies what it is for logic to be relevant). The second is to augment the language of the original consequence relation with a conditional which mirrors the relevant consequence relation.[302]

This is not the place to discuss the details of the construction, so I will just note a couple of points. The strategy of starting with a consequence relation and filtering out the irrelevant instances is a well known one. The logics produced are often termed 'filter logics' for obvious reasons.[303] Typically, though not invariably, filter logics are non-transitive, as is Verdée's logic. Secondly, the strategy employed by Verdée is very close to that which generates Tennant's Core Logic.[304] Both function by taking a consequence relation and filtering out those inferences where a premise or conclusion is redundant.[305] Given that the classical inference $A, \neg A \vdash B$ already falls by the wayside on this approach, one might wonder what the benefits are of moving away from classical logic to apply the strategy. I will leave that matter for

[300]Thus, one can formulate the sorites paradox with different sorts of conditional, and different solutions may be appropriate for different conditionals. For some it may be appropriate to say that the conditional premises are untrue; for some it may be appropriate to say that detachment is invalid. (See Priest (2010b), §7.)

[301]See Priest (2014), §15.

[302]On a small matter: Verdée uses the word 'implication' for both a conditional connective and a validity relation. Of course, he is not confused about the distinction, and there is a venerable tradition using the word for the conditional. However, I think that the terminology is unfortunate. 'Implies' is not a connective; it is a relation. And in my experience, the terminology encourages a confusion in students between the connective and the consequence relation.

[303]For a discussion and references, see Priest (2002), §4.1 .

[304]See §23 above.

[305]The easiest way to see this is to compare Verdée's Definition 5 and Theorem 1 with the the presentation of Tennant's approach in Priest (2002), §4.1.

Verdée to address.

Finally, I note that Verdée's conditional is distinct from that of all the usual relevant logics. It is not contained in any of these, since it does not contain transitivity. And it does not contain any of these, since, as Verdée notes (§5), it verifies things such as $A \rightarrow (B \rightarrow (A \wedge B)$, which does not hold in the usual systems of relevant logic. So there is an issue as to which of the two approaches gives the best sort relevant conditional—however one might understand that question. That, however, is too big an issue to take on here.

25 Wansing and Skurt: If a is b, and b is c, a is c, isn't it?

My training in logic was very much in classical mathematical logic. I suppose that as the years have gone on, I have come to see how the mathematical techniques I learned could be applied to produce many other logics—non-classical logics. These, in turn, opened up new avenues of approach to philosophical problems, both traditional and contemporary. This is very much true concerning the topic of identity. I started by assuming, like most contemporary logicians, that the standard text-book account of identity gets it right. Nothing made me challenge this assumption for many years. However, I started to think otherwise in the 1990s, when I came to realise how an account according to which the transitivity of identity (TI) fails could be applied to solve a number of interesting philosophical problems, and how, plausibly, my thinking had been channelled by what Wittgenstein called an inadequate diet of examples (mainly from mathematics).[306] The first ideas in this direction concerned the failure of transitivity in continuum-valued logics, but in ONE the failure of transitivity was obtained by defining $x = y$ as $\mathfrak{A}P(Px \equiv Py)$ in second-order LP.

25.1 Identity Itself

In their paper, Heinrich Wansing and Daniel Skurt (hereafter W&S)—whom I came to know very well in a brief but happy sojourn at the Ruhr University

[306] 'A main cause of philosophical disease—a one-sided diet: one nourishes one's thinking with only one kind of example'. *Philosophical Investigations,* §593. Wittgenstein (1968), p. 144^e.

of Bochum in 2013—take us into this world of non-transitive identity. They make a number of interesting technical observations. However, in keeping with my policy in writing these comments, I shall not go into technical matters. What I will do is to consider the objections they raise to the account of identity in ONE—which, to eliminate any uncertainly about my view (of the kind that W&S find in their §6), is my current view.[307]

First, W&S insist on calling identity according to the standard account *real* identity (e.g. §§3, 7). Indeed, even the title of their essay uses scare quotes when referring to non-transitive identity.[308] Now that is an entirely tendentious way of putting the matter. I take it that, orthodox as the account may be, it in fact mis-characterises identity—though the characterising conditions may well hold in "normal" circumstances. Whether the standard account of identity gets the real identity relation right is exactly what is at issue in this matter.[309]

It remains the case that one may, in the metalanguage, have a notion of identity, $\asymp$, specified as in the standard account, and so where its extension is $\{\langle x, x\rangle : x \in \mathcal{D}_1\}$, $\mathcal{D}_1$ being the first-order domain).[310] However, it does not follow that this should be used to state the truth conditions of $=$ in the object language. Why not? Because it is not identity (or at least one can assume so only by begging the question)! What, then is the relation $\asymp$? The simple answer is that it is co-substitutivity, a relation that is, of course, language-dependent.

W&S's §3 raises the question of similarity relations. Similarly relations are reflexive and symmetric, but not transitive. So it may occur to the reader to wonder why $=$, as I have defined it, is no more than a similarity relation. The answer is simple. A similarity relation may not satisfy the condition $\mathfrak{A}P(Px \equiv Py)$. Consider being similar with respect to colour; and let a and b be distinct objects that are similar in this respect. a, we may suppose, emits light of frequency ν_a, and b emits light of frequency ν_b. Let Q be the

[307] I note, as they do (§6), that the non-transitivity of the identity relation and its inconsistency are, in principle, quite different issues. There are theories of identity which verify each of these but not the other.

[308] They say 'non-transitive "identity" is not identical with what is usually taken to be real identity' (§7). That, at least, is true. But note, again, the scare quotes around 'identity'.

[309] In §2 they quote me as saying that TI for real identity fails. But as I would hope the context makes clear, by 'real identity' I am referring to identity as it really is, and not how some mistaken theory takes it to be.

[310] The following points are made in ONE, 5.11.

property of emitting light of frequency ν_a (and suppose this to be a consistent condition). Then Qa is true (and true only) while Qb is false (and false only). It follows that $\mathfrak{A}P(Pa \equiv Pa)$ is not true.

25.2 Second-Order *LP*

§4 of W&S's paper raises questions concerning the range of the second-order quantifiers, second-order minimal inconsistency, and reassurance.

Let $\mathcal{D}_2$ be the second-order domain. If this contains every extension/anti-extension pair of the form $\langle X, Y\rangle$ such that $X \cup Y = \mathcal{D}_1$—call this the *full* $\mathcal{D}_2$—or even just the pair $\langle \mathcal{D}_1, \mathcal{D}_1\rangle$, then every sentence of the form $a = b$ is true and false. As W&S note, this is an undesirable consequence. But as W&S also note, ONE puts no constraints on $\mathcal{D}_2$ other than that it be non-empty. I left it open what other constraints $\mathcal{D}_2$ should satisfy, though I was very clear (2.7) that it should not be full, to rule out exactly the sort of consequences that W&S point to. The denizens of the domain need to be metaphysically real properties.[311]

Next, W&S raise the question of classical recapture with respect to second-order minimal *LP*, *LPm*. Classical recapture is a statement to the effect that for consistent Σ, if A follows from Σ in classical logic, it follows in *LPm*. In their definition (Definition 2) of the consistency ordering, $\prec$, they impose a constraint to the effect that if $\mathcal{I} \prec \mathcal{J}$ then the first-order domain of $\mathcal{I}$ is a superset of the first-order domain of $\mathcal{J}$.[312] They then note that classical recapture is problematic. Perhaps so. However in the discussion of first-order *LPm* in IC, 16.4-16.5, no such constraint is imposed; and that being so, classical recapture is immediate. Neither is it imposed in the discussion of the second-order case in ONE, 5.13. And as I note there, this gives recapture for classical first-order logic *with* (standard) *identity*; though what happens

[311]In their paper, W&S refer to Hazen and Pelletier (2018), who prove a number of interesting results about second-order *LP* and the Leibniz-defined identity in this. Hazen and Pelletier refer in their Abstract to the characterisation of $\mathcal{D}_2$ (if I understand the allusion correctly), saying that 'it will be extremely difficult to appeal to second-order *LP* for the purposes that its proponents advocate, until some deep, intricate, and hitherto unarticulated metaphysical advances are made'. They surely overplay their hand here. One does not have to determine every parameter of a piece of logical machinery before one can reasonably deploy it—especially since the applications can provide a constraint on fixing those parameters.

[312]They go on to call this clause 'unintuitive'. My intuitions—for or against—do not stretch this far.

with second-order inferences may depend on how the details of the second-order domain are filled out. The first-order inferences would seem to be good enough for most practical purposes, however. After all, it's all you have on the standard account favoured by W&S.

So much for classical recapture. Reassurance is another matter. Reassurance is the claim that if Σ is non-trivial under LP, it is non-trivial under LPm. This holds for propositional logic, but as W&S note, extending the result to first-order logic—and *a fortiori* second-order logic—has turned out to be a very tricky matter,[313] and domain-restriction conditions play an important role in the matter. Much work remains to be done to sort things out. Here I note only two things.

First, at one time I did take reassurance to be a necessary condition for a suitable notion of minimal inconsistency. I no longer think so. If it holds "for the most part", that will be fine. If I may quote myself from Priest (2017), §6.3:

> Now, even without Reassurance, LPm would seem to do everything that one would like: it delivers a more generous notion of consequence than LP, where irrelevant contradictions do not invalidate classical inferences, and which delivers all classical consequences given consistent premises. [Footnote: The reason given in IC, 16.6, for the desirability of Reassurance is as follows. Taking triviality to be a mark of incoherence, Reassurance guarantees that a coherent situation will never be turned into an incoherent one under LPm. This may be more than is required, though. It might be quite sufficient if mostly, or normally, LPm does not turn a non-trivial situation into a trivial one. If there are some exceptions, and LPm is otherwise robust, we might take the triviality exposed to speak against the coherence of the original situation.]

As I go on to note there, there is, in any case, a very easy way to obtain reassurance. One simply redefines the LPm consequence relation as follows. $\Sigma \vDash_m A$ iff:

- (some minimally inconsistent model of Σ is non-trivial, and every minimally inconsistent model of Σ is a model of A) or (every minimally

[313] As Crabbé (2011) and (2012) have made clear.

inconsistent model of Σ is trivial—in particular, if there are none—and $\Sigma \vDash_{LP} A$).

Reassurance follows simply, and $\vDash_m$ still has all the crucial properties.[314]

At the end of the section, W&S state that the open questions they note on these matters 'need to be answered before one can seriously consider using [Priest's] definition of identity in, for example, paraconsistent set theory or for solving paradoxes like the sorites paradox'. To the extent that these questions need answering, I take it that I have now answered them.

25.3 The Leibniz Definition

Finally, in §6 W&S turn to the matter of whether identity should be defined *via* the Leibniz condition. They quote a passage from Williamson (2006) saying that:

> it is unlikely that second-order quantification is conceptually more basic than identity in any deep sense. But the appeal to second-order quantification may not satisfy those who are seriously worried about the problem of interpreting the identity predicate. For how do we know, or what makes it the case, that the second-order quantifier $\forall P$ should be interpreted in the standard way?

The standard way referred to in the last sentence is where the second-order domain is the full power set of the first order-domain. (Williamson's discussion is in the context of classical logic.)

We may ignore the point about being conceptually more basic. If people subscribe to some kind of methodology of conceptual foundationalism, I leave it to them to figure out what is more fundamental than what. For my part, I take all such programs of conceptual analysis to be flawed.

As for the second point, the context of Williamson's quotation is a concern with how one can be sure that someone who characterises identity by the usual first-order conditions really means identity (as usually understood), since there are non-standard interpretations of the language in which it is not so interpreted. He is pointing out that appealing to a characterisation of identity by the Leibniz condition in second-order logic is of no help here, since there are non-standard interpretations of the second-order quantifiers

[314] As I also note there, Batens works with a different notion of minimal inconsistency, which guarantees Reassurance.

as well. Now, what, exactly, the intended interpretation of the second-order logic is in the present case is as yet undetermined, pending a specification of the range of the second-order variables. But in any case, Williamson's comments are beside the point in the present context. True, a syntactic characterisation of the second-order semantics in question is likely to underdetermine; but as the context makes clear, this is equally the case with the standard first-order theory of identity. As we know, any axiomatic theory in the language of first- or or second-order logic with an infinite model will have non-standard interpretations.[315] The non-standard are always with us. What, then, about usage ensures that we have a standard interpretation is an important question; but it is not one on the agenda here.

W&S then refer to a passage from Manzano (2005). In the context, she is addressing the question of why, though one may define standard identity in (full) classical second-order logic, it is a good idea to take it as primitive anyway. The first three reasons she gives are to the effect that in less than full second-order logic the Leibniz definition may not deliver standard first-order identity. As hardly needs to be said, this is not a reason for not characterizing identity with the Leibniz condition if the aim is not to recapture the usual account of identity—she even points out that the substitutivity of identicals may fail!—and, moreover, one is not concerned with subsystems of classical logic anyway. Such considerations are simply, therefore, beside the point.[316]

In the last paragraph of the section W&S, cite two further problems they see with defining identity by the Leibniz condition in *LP*. The first is the failure of *modus ponens* for the material conditional in *LP*. As the saying goes, this is not a bug, it is a feature. It is this which undergirds the account of unity in ONE! It might be suggested that since most contemporary logicians hold identity to be transitive, its failure requires some "independent motivation"; and perhaps the gluon theory of ONE is unusual enough not to count. But as ONE, 5.7, 5.8 notes, there are well-known (if frequently ignored) apparent counter-examples to TI. And as 5.3 points out, in this context, standard arguments for TI are bankrupt.

[315]Actually, in the case of *LP*, the restriction to having an infinite model can be dropped, due to the Collapsing Lemma.

[316]Manzano's fourth point is slightly different. This is that (classical) primitive identity, comprehension, and extensionality allow us to introduce set/property abstracts. Outwith this context, abstracts have to be taken as primitive. I see no reason why this is problematic. If abstracts are required, it seems no worse to take these as primitive than to take identity itself as primitive.

Of course, it might be suggested that the LP biconditional is the wrong one to use in the Leibniz condition, and that a detachable conditional should be used. That matter is taken up in ONE, 2.6.

The second of W&S's supposed problems is that the possibility of contradictory identities is inherited from other contradictions. I simply fail to see this as a problem. Why should the contradictory nature of identity not follow from other contradictions? As ONE, ch. 1, argues, one should expect gluons to be inconsistent objects. One is, hence, committed to a paraconsistent logic. As explained in ch. 2, the Leibniz definition of identity then gives rise to a non-transitive identity. This explains *why* identity is not transitive, when one might have expected otherwise. It also explains why gluons do not generate a Bradley-style regress, solving the problem of unity. Dialetheism about gluons is, therefore, fundamental to the whole story.

In §7 W&S add a final complaint:

> [Priest defines] identity of individuals as Leibniz-identity in second-order minimal LP. The choice, though clearly motivated, is not unproblematic, in particular if it is driven to the point that second-order minimal LP is to be used as a metatheory, thereby expulsing the standard notion of real identity entirely.

Now, as I have explained in 3.2 above, the semantic metatheory is best thought of in terms of ZF, which one can make sense of in LP-based (not, *nota bene*, LPm-based) naive set theory. And matters concerning identity are irrelevant. For as is well known, ZF can be expressed in a language with a single predicate, $\in$, the principle of extensionality being formulated as $\forall z(x \in z \equiv x \in y) \supset (x \in w \equiv y \in w)$.

In their final sentence W&S say that they prefer the standard theory of identity to my account. They are, of course, entitled to their preferences. However, they do not say how they would solve the philosophical problems that a non-transitive identity allows. Nor, as I have indicated, do I find their problems with non-transitive identity very persuasive.

26 Weber: Contradictions Before the Limit

Weber's results on relevant naive set theory are clearly the most significant advances in the topic since, and on a par with, Brady's groundbreaking work in the late 1970s. Brady showed what *could not* be proved: suitably formulated, relevant naive set theory was non-trivial. The other side of the coin is what *could* be proved. Could one establish standard results of the theory of transfinite cardinals and ordinals in the theory? Many people in the Canberra group of the late 1970s (including myself) struggled with this problem, and gave up. Weber showed how it could be done some 30 years later.[317]

Over recent years, Weber has also been the most ardent advocate of dialetheism, as well as one who has significantly stretched the dialetheic frontiers. In his essay in this collection, he returns to what first attracted him to dialetheism: the Inclosure Schema (IS). There is much I agree with in the essay, and some things with which I do not. I think the best way to put the matter into focus is, again, to provide an historical perspective.

IC was a full frontal assault on the Principle of Non-Contradiction. The paradoxes of self-reference loomed large in this. I used them as a battering ram, as it were, on what I took to be the weakest part of the defences, eroded, as they had been for decades—indeed, for centuries, for those who know their history of Western logic—by failure to reach consensus. But these were not the only applications of dialetheism there, as the third part of the first edition makes clear.[318]

Thanks to discussions with Uwe Petersen in the mid 1980s I came to see the connection between the paradoxes of self-reference and the limit phenomena important to the thought of Kant and Hegel. The result was BLoT. The main thesis of this was that there are certain limits that are dialetheic (the limit of what can be expressed; the limit of what can be described or conceived; the limit of what can be known; the limit of iteration of some operation or other, the infinite in its mathematical sense[319]). These are all such that one can go no further; yet one can.

In the process of writing BLoT I formulated the Inclosure Schema. An inclosure arises where there is a totality, Ω, and an operator, δ, such that it appears to be the case that when δ is applied to any subcollection of Ω of an appropriate kind, it produces an object not in that subcollection (Transcendence) but still in Ω (Closure). These conditions obviously produce contradiction when applying δ to Ω itself. I took the idea of the IS essentially from Russell, though I tweaked it to accommodated paradoxes such as the

[317] This was in his doctoral thesis, much of which appeared in Weber (2010) and (2012).

[318] I give a category of some possible applications of dialetheism in §11.1 above. 1-5 and 8 all appear in IC1.

[319] BLoT, p. 3.

Liar cleanly.

The IS was never meant to be an argument for dialetheism; that was not its point[320]—though careless expression sometimes might have suggested otherwise. It was a diagnostic schema for characterising a class of paradoxes (hence unifying them), and showing why they arise. What to do about them is another matter. Russell, after all, was no dialetheist. Of course, if the conditions of the IS are true, then so is the contradiction they generate; and the dialetheic solution is simply to accept this.

The main application of the IS in BLoT is as an argument against orthodoxy (of the time and contemporary) which espouses quite different solutions to the set-theoretic paradoxes and the semantic paradoxes. These both fit the IS, and so, being the same kind of paradox, should have the same kind of solution (the Principle of Uniform Solution).[321]

It had always seemed to me that sorites paradoxes were of a quite different kind from paradoxes of self-reference, and I was not at all inclined to a dialetheic solution for them. But in discussion with Mark Colyvan in 2007, it struck me that the sorites paradoxes were inclosure paradoxes (that is, paradoxes fitting the IS) as well. Because of the Principle of Uniform Solution, I therefore came to accept a dialetheic solution.

26.1 Limits and Dialetheism

With these preliminary considerations, we can now turn to Weber's paper.

First, and most straightforwardly, Weber shows (§§4.2, 5) that, on pain of triviality, not all limits can be transcended. True; but the aim of BLoT was not to show that they could be.[322] (I don't think it ever occurred to me that this might be the case.) So to the extent that Weber interpreted me as saying so (fn. 22), that *is* a misinterpretation.[323] To transcend a boundary, there must be something that takes you to the other side. Such, though contradictory, is the case for the limits of thought with which BLoT deals. The limit of the non-trivial rules this out—on pain of triviality.

[320] As Weber notes at the end of §3.1.

[321] BLoT, 11.5.

[322] 'Limits of *this kind* provide boundaries beyond which certain processes... cannot go; a sort of conceptual *ne plus ultra*. The thesis of this book is that such limits are dialetheic...' BLoT, p. 3. (First italics added here.)

[323] Weber also quotes me as quoting Cantor talking about breaking through every barrier. The 'every' here concerns just ordinals.

The next topic (the rest of §4) concerns problems Weber finds with the IS itself. He makes two points in this context. The first (§4.1) is that there are arguments that have the right shape for an inclosure, but are not paradoxes. One example concerns an omnipotent god; a second concerns diagonalising out of the natural numbers.[324] Some have made a similar point using the "barber paradox".[325]

As Weber notes (and BLoT 17.2 explains) for the IS conditions to create paradoxes we have to have good reasons, of at least a *prima facie* kind, to suppose that the conditions are true. There are none in these cases. Weber concurs, but objects:

> Prima facie validity will vary from reasoner to reasoner, though, depending on their logical training. And more importantly, 'seeming' validity is not a reliable guide to truth. Once one gets past initial diagnostic devices, more precise and reliable instruments than mere prima facie plausibility are needed. The way to distinguish a genuine contradiction from a contradiction-shaped joke is to show that the conditions of the inclosure are genuinely satisfied.

Now the first point is true to a certain extent, but should not be over-played. The naive principles that generate the paradoxes of self-reference, and the reasoning from them, have appeared *prima facie* obvious to all those who have thought about them. That is why, after all, we call these things paradoxes, and not simply *reductio* arguments.

The rest of Weber's quote is also agreed; but the IS was never meant to be an argument for dialetheism, as I have explained. It just characterises a class of paradoxes. It is an independent question as to what to do about them. For the contradictions they generate to be true, the IS conditions have to be true. That will require separate arguments. A major one such is that no consistent solution to the paradoxes works, simply because the principles it invokes merely succeed in relocating the inclosure.[326]

[324] Actually, I don't how the first of these is supposed to fit into the IS. What, for example, is the diagonaliser? But that is of no great moment here.

[325] See BLoT2, 17.2

[326] BLoT, pp 228 ff.

26.2 Reasoning to the IS Conditions

Weber's next point is more substantial. The IS conditions are generated by naive principles about truth, sethood, and related matters. But to establish the IS conditions we need to reason from these principles. Weber points out that some of the reasoning involved may not be dialetheically valid. Point, again, taken.[327] In BLoT (p. 130, fn 7) drawing on the fact that if Σ entails A in classical logic then, for some B, Σ entails $A \vee B!$ in LP (where $B!$ is $B \wedge \neg B$), I noted that even if we cannot establish that $\delta(\Omega) \in \Omega!$ in a kosher fashion, we still have a contradiction in the form $\delta(\Omega) \in \Omega! \vee B!$. Thus, the IS conditions still deliver contradiction.

Weber notes this, but observes, correctly, that, in this case, the contradiction may no longer concern the limit (Ω). True; though most of the arguments for instances of the IS conditions given in BLoT *are* dialethically valid (for example, those concerning truth, knowledge, expressibility), as may be checked on a case by case basis.

Weber analyses a couple of cases where they are not. It will pay to look at these more closely. One is a version of Russell's paradox where Ω is V (§4.2.1). He points out that in proving Transcendence, given $X \subseteq \Omega$, and having shown that $\delta(X) \notin \delta(X)$, the best we can do is to establish that:

- $\delta(X) \in X \wedge \delta(X) \notin \delta(X) \vdash \delta(X) \in \delta(X)!$

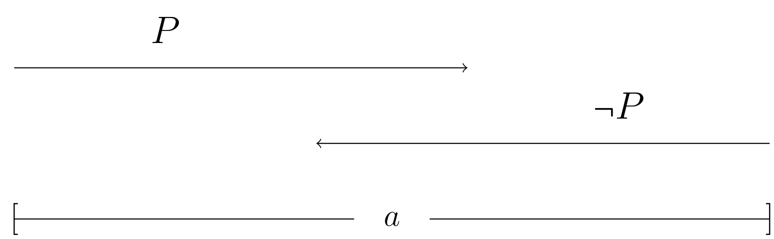
(changing his r to δ). From this we cannot get $\delta(X) \notin X$ by valid means. But given that $\delta(X) \notin X$ or $\delta(X) \in X$, we have either $\delta(X) \notin X$ or $\delta(X) \in \delta(X)!$, so if transcendence fails, this is because there are already inconsistent boundaries (within Ω), as BLoT claims

The second case that Weber analyses is the sorites paradox (which does not feature in BLoT). Given a sorites sequence of objects, $\Omega = \{a_0, ..., a_n\}$ is the set of things satisfying some vague predicate, P; and if $X \subseteq \Omega$, $\delta(X)$ is a_{i+1}, where i is the largest j such that $a_j \in X$. As Weber points out (§4.2.4), given that the conditionals in the sorites argument are material, one cannot establish Closure. If $X \subseteq \Omega$, it needs to be shown that Pa_{i+1}. The closest

[327] I agree with Weber that the logic of an acceptable metatheory should be paraconsistent, though, despite what Weber says, the point appears to me to be irrelevant here. For the arguments to work, they have to be valid; one not does not have to prove them to be so in a metatheory. For a discussion of paraconsistent metatheory, see the reply to Batens, 3.2 above.

we can get is $Pa_i! \vee Pa_{i+1}$. What the conditions of the sorites entail is that $\bigvee_{0\leq i\leq n} Pa_i!$.[328] This does not show where the contradiction lies.

However, if one consults the dialetheic models of a sorites transition, in which borderline elements are contradictory, we have something of the form:[329]



Hence, if a is in the borderline area, $Pa_i!$ holds. There is a clear sense in which such as constitute the boundary of Ω. Hence, again, the boundary is indeed contradictory. Similarly, such as constitute the boundary of $\overline{\Omega} = \{x : \neg Px\}$, so the boundary of this is also contradictory.[330]

26.3 Local Contradictions

Having said these things, it remains the case, as Weber points out, that inclosures may be within a larger totality. He establishes this with respect to the IS and another version of Russell's paradox, where Ω is the Russell set, R, itself (not V). As demonstrated, this is an inclosure contradiction. However, he says that:

> whether this is a meaningful instance of the inclosure schema is dubious. The diagonalizer is doing nothing; all the energy is coming from the totality, top down, rather than a contradiction surging towards the totality.

I fail to see this. The diagonaser is still transcending lesser collections. It's just that the eruption occurs only part-way up the mountain (of the absolute), to pick up the metaphor. Note also that the domain of the inclosure in this

[328] Priest (2010b), §5.
[329] Priest (2010b), §4.
[330] This is an example of the "double inclosure" in Weber's appendix

case is R, so the contradiction does occur at the boundary of the relevant inclosure.

More generally, it was never an aim of BLoT to show that dialetheias occur *only* at absolute infinities. Even in BLoT, there are many inclosure contradictions which are not of this kind: the liar (where Ω is the set of truths; Berry's paradox, where Ω is a subset of the natural numbers; König's paradox, where Ω is the set of definable ordinals). These contradictions arise a long way from absolute infinity, and so are hardly contradictions at 'the edge of the universe' (as Weber puts it §6). And as BLoT, pp. 170 f. shows, there can indeed be inclosures within inclosures.

More generally, as I have noted, IC argues for the existence of many "local" dialetheias—concerning motion and law, for example. These are not only nothing to do with absolute infinity; they are not even inclosure contradictions. So I quite agree with Weber's comment (§1) that 'inclosure-based dialetheism is not enough'.

IC, ch. 8, argues that since dialetheias are relatively rare, we may use classical logic as a default inference-engine until one is shown that one cannot.[331] Weber worries (§6) that this rash of local contradictions will make the methodology vacuous. I don't think so. The question is how to understand rarity. Even if there is just one dialetheia, p_0, there are as many dialetheias as there are formulas. (If q is any true statement, $p_0 \wedge q$ is another dialetheia.)[332] The point is that dialetheias are statistically infrequent in our reasoning.[333] Grant that dialetheias occur at instantaneous states of change. We rarely reason about the truly instantaneous states. Grant that dialetheias occur in the border-areas of vague predicates. We rarely reason about such areas. (Even if the predicate 'red' is vague, most things are simply not red.) Similar remarks apply to paradoxical sets and truths.

To summarise my agreements and disagreements with Weber: Dialetheias are many. Some of these are inclosure contradictors; some are not. And even where they are inclosure contradictions, the domain of the inclosure may not be absolutely infinite, but some much smaller totality. It remains the case that inclosure contradictions show there are limits of certain notions (the

[331] I agree with Weber (§6) that 'deductive reasoning should never presume consistency'. The inference engine is one of default reasoning, and so non-deductive, as shown most clearly when the strategy is implemented in LPm.

[332] So, again, I agree with Weber (§6) that all objects will satisfy inconsistent conditions. Let a be any object, and P any condition that it satisfies; then $p_0 \wedge Pa$ and $\neg(p_0 \wedge Pa)$.

[333] See IC, 8.4.

true, the expressible, and so on), where the notions go no further... though they do. All as BLoT says.

27 Conclusion: Looking Back Over My Shoulder

As I have been writing these comments over the last six months, I have had to go back and read—or reread—a number of things I have written on matters paraconsistent and dialetheic over the last 40 years or more. In doing so, I have come face to face with the way in which my own thinking on these matters has evolved during the course of that journey. I was struck by how impossible it would have been to predict how things would evolve—and continue to evolve: in writing these comments I have had to think through a number of matters afresh.

Be that as it may, the journey I have made, I have not made on my own. Friends, colleagues, students, and many whom I hardly know—some now, sadly, dead—have been fellow-travellers. Their investigations and critical insights, positive and negative, have all helped to enrich the road. I am indeed fortunate to have been accompanied by such good friends.

References

[1] Alxatib, S., and Pelletier, F. (2011), 'The Psychology of Vagueness: Borderline Cases and Contradictions', *Mind and Language* 26: 287-326.

[2] Beall, J. C. (2009), *Spandrels of Truth*, Oxford: Oxford University Press.

[3] Beall, J. C. (2011), 'Multiple-Conclusion *LP* and Default Classicality', *Review of Symbolic Logic* 4: 326-36.

[4] Beall, J. C. (2012), 'Why Priest's Reassurance is not so Reassuring', *Analysis* 72: 517-25.

[5] Beall, J. C., and Armour-Garb, B. (eds.) (2005), *Truth and Deflationism*, Oxford: Oxford University Press.

[6] Beall, J. C., Brady, R., Dunn, J. M., Hazen, A., Mares, E., Meyer, R., Priest, G., Restall, G., Ripley, D., Slaney, J., and Sylvan, R. (2012), 'On the Ternary Relation and Conditionality', *Journal of Philosophical Logic* 41: 595–612.

[7] Beall, J. C., Brady, R., Hazen, A., Priest, G., and Restall, G. (2006), 'Relevant Restricted Quantification', *Journal of Philosophical Logic* 35: 587-98.

[8] Bell, J., Cole, J., Priest, G., and Slomson, A. (1973), *Proceedings of the Bertrand Russell Memorial Logic Conference*, Leeds: University of Leeds.

[9] Belnap, N. (1977), 'A Useful Four-Valued Logic', pp. 5-37 of J. M. Dunn and G. Epstein (eds.), *Modern Uses of Multiple-Valued Logic*, Dordrecht: Reidel.

[10] Berto, F., French, R., Priest, G., and Ripley, D. (201+), 'Williamson on Couterpossibles', to appear.

[11] Brady, R. (1989), 'The Non-Triviality of Dialectical Set Theory', ch. 16 of Priest, Routley, and Norman (1989).

[12] Brown, B. (1999), 'Yes, Virginia, There Really Are Paraconsistent Logics', *Journal of Philosophical Logic* 28: 489-500.

[13] Brown, B., and Priest, G. (2004), 'Chunk and Permeate I: the Infinitesimal Calculus', *Journal of Philosophical Logic* 33: 379-88.

[14] Caie, M. (2012), 'Belief and Indeterminacy', *Philosophical Review* 121: 1-54.

[15] Cantini, A. (2003), The Undecidability of Grišin's set theory', *Studia Logica*, 74: 345-368.

[16] Carnielli, W., and Rodrigues, A. (2017), 'An Epistemic Approach to Paraconsistency: a Logic of Evidence and Truth', *Synthese*, doi: 751 https://doi.org/10.1007/s11229-017-1621-7.

[17] Castañeda, H. N. (1989), *Thinking, Language, and Experience*, Minneapolis, MN: University of Minnesota Press.

[18] Chalmers. A. (2014), *What is This Thing Called Science?*, 4th edn, Brisbane: University of Queensland Press.

[19] Church, A. (1936), 'An Unsolvable Problem of Elementary Number Theory', *American Journal of Mathematics* 58: 345-363.

[20] Church, A. (1956), *Introduction to Mathematical Logic*, Princeton, NJ: Princeton University Press.

[21] Cobreros, P., Égré, P., Ripley, D., van Rooij, R (2012), 'Tolerant, Classical, Strict', *Journal of Philosophical Logic* 41: 347-85.

[22] Cobreros, P., Égré, P., Ripley, D., van Rooij, R (2013), 'Reaching Transparent Truth', *Mind* 122: 841-66.

[23] Cotnoir, A., and Weber, Z. (2015), 'Inconsistent Boundaries', *Synthese* 192: 1267-94.

[24] Crabbé, M. (2011), 'Reassurance for the Logic of Paradox', *Review of Symbolic Logic* 4: 479-85.

[25] Crabbé, M. (2012), 'Reassurance *via* Translation', *Logique et Analyse* 55: 281-93.

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

[27] Dummett, M. (1978a), 'The Philosophical Significance of Gödel's Theorem', ch. 12 of *Truth and Other Enigmas*, London: Duckworth.

[28] Dummett, M. (1978b), 'The Philosophical Basis of Intuitionist Logic', ch. 14 of *Truth and Other Enigmas*, London: Duckworth.

[29] Dunn, J. M. (1979), 'A Theorem in 3-Valued Model Theory, with Connections to Number Theory, Type Theory, and Relevant Logic', *Studia Logica* 38: 149-69.

[30] Field, H. (1973), 'Theory Change and Indeterminate Reference', *Journal of Philosophy* 70: 462-81.

[31] Field, H. (2005), 'Is the Liar Both True and False?, ch. 2 of Beall and Armour-Garb (2005).

[32] Field, H. (2008), *Saving Truth from Paradox*, Oxford: Oxford University Press.

[33] Field, H. (2014), 'Naive Truth and Restricted Quantification', *Review of Symbolic Logic* 7: 147-91.

[34] Field, H. (2016), 'Indicative Conditionals, Restricted Quantification and Naive Truth', *Review of Symbolic Logic* 9: 181-208.

[35] Field, H. (201+), 'Properties, Propositions and Conditionals', *Australian Philosophical Review*, forthcoming.

[36] Fine, K. (1975), 'Vagueness, Truth, and Logic', *Synthese* 30: 265-300.

[37] Glanzberg. M. (2013), 'Truth', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/truth/.

[38] Goodship, L. (1996), 'On Dialethism', *Australasian Journal of Philosophy* 74: 153-61.

[39] Harman, G. (1986), *Changes in View*, Cambridge, MA: MIT Press.

[40] Hazen, A., and Pelletier, J. (2018), 'Second-Order Logic of Paradox', *Notre Dame Journal of Formal Logic* 59: 547-58.

[41] Humberstoneone, L. (2011), *The Connectives*, Cambridge, MA: MIT Press.

[42] Keefe, R., and Smith, P. (eds.) (1999), *Vagueness: a Reader*, Boston: MIT Press.

[43] Kripke, S. (1975), 'Outline of a Theory of Truth', *Journal of Philosophy* 72: 690-716.

[44] Lakatos, I. (1978), 'Cauchy and the Continuum: the Significance of Non-Standard Analysis for the History and Philosophy of Mathematics', *The Mathematical Intelligencer* 1: 151-61; reprinted as pp. 43-60, Vol. 2, of Lakatos' *Collected Papers*, eds. J. Worrall and G. Currie, Cambridge: Cambridge University Press, 1978.

[45] MacFarlane, J. (2004), 'In What Sense (if any) is Logic Normative?', a talk presented to the Central Division of the American Philosophical Association.

[46] Manzano, M. (2005), *Extensions of First-Order Logic*, Cambridge: Cambridge University Press.

[47] Martin, B. (2014), 'Dialetheism and the Impossibility of the World', *Australasian Journal of Philosophy* 93: 61-75.

[48] Mares, E. (1996), 'Relevant Logic and the Theory of Information', *Synthese* 109: 345-60.

[49] Mares, E. (2004), *Relevant Logic: a Philosophical Interpretation*, Cambridge: Cambridge University Press.

[50] Maund, J. B. (1981), 'Colour: a Case of Conceptual Fission', *Australasian Journal of Philosophy* 59: 308-22.

[51] Nelson, E. (2015), 'Inconsistency of Primitive Recursive Arithmetic', *arXiv*, https://arxiv.org/abs/1509.09209.

[52] Omori, H. (201+), 'From Logics of Formal Inconsistency to Logics of Formal Classicality', *Logic Journal of the IGPL*, forthcoming.

[53] Paris, J., and Sirokofskich, A. (2008), 'On *LP*-Models of Arithmetic', *Journal of Symbolic Logic* 73: 212-226.

[54] Priest, G. (1979), 'The Logic of Paradox', *Journal of Philosophical Logic* 8: 219-41.

[55] Priest, G. (1984), 'Logic of Paradox Revisited', *Journal of Philosophical Logic* 13: 159-73.

[56] Priest, G. (1987), *In Contradiction*, Dordrecht: Martinus Nijhoff; 2nd edn, Oxford: Oxford University Press, 2006.

[57] Priest, G. (1990), 'Dialectic and Dialetheic', *Science and Society* 53: 388-415.

[58] Priest, G. (1991), 'Minimally Inconsistent *LP*', *Studia Logica* 50: 321-31.

[59] Priest, G. (1994a), 'Is Arithmetic Consistent?', *Mind* 103: 337-49.

[60] Priest, G. (1994b), 'What Could the Least Inconsistent Number Be?', *Logique et Analyse* 37: 3-12.

[61] Priest G. (1995a), *Beyond the Limits of Thought*, Cambrige: Cambridge University Press; 2nd edn, Oxford: Oxford University Press, 2002.

[62] Priest, G. (1995b), 'Multiple Denotation, Ambiguity, and the Strange Case of the Missing Amoeba', *Logique et Analyse*, 38: 361-73.

[63] Priest, G. (1998), 'What's so Bad about Contradictions?', *Journal of Philosophy* 95: 410-26; reprinted as ch. 1 of G. Priest, J. C. Beall, and B. Armour-Garb, *New Essays on the Law of Non-Contradiction*, Oxford: Oxford University Press.

[64] Priest, G. (1999), 'Validity', pp. 183-206 of A. Varzi (ed.), *The Nature of Logic* (*European Review of Philosophy*, vol. 4), Stanford, CA: CSLI Publications; reprinted as ch. 11 of Priest (2006a).

[65] Priest, G. (2001), 'Logic: One or Many?', pp. 23-38 of J. Woods and B. Brown (eds.), *Logical Consequences: Rival Approaches*, Oxford: Hermes Scientific Publishers; reprinted as ch. 12 of Priest (2006a).

[66] Priest, G. (2002), 'Paraconsistent Logic', pp. 287-393, Vol. 6 of D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, 2nd edition, Dordrecht: Kluwer Academic Publishers.

[67] Priest, G. (2003), 'Inconsistent Arithmetic: Issues Technical and Philosophical', pp. 273-99 of V. F. Hendricks and J. Malinowski (eds.), *Trends in Logic: 50 Years of Studia Logica*, Dordrecht: Kluwer Academic Publishers.

[68] Priest, G. (2005a), *Towards Non-Being*, Oxford: Oxford University Press; 2nd edn, 2016.

[69] Priest, G. (2005b), 'Spiking the Field Artillery', ch. 3 of Beall and Armour-Garb (2005).

[70] Priest, G. (2006a), *Doubt Truth to be a Liar*, Oxford: Oxford University Press.

[71] Priest, G. (2006b), 'The Paradoxes of Denotation', ch. 7 of T. Bolander, V. F. Hendricks, and S. A Pedersen (eds.), *Self-Reference*, Stanford, CA: CLSI Lecture Notes.

[72] Priest, G. (2007), 'Reply to Slater', pp. 467-74 of J. Y. Béziau, W. Carnielli, and D. Gabbay (eds.), *Handbook of Paraconsistency*, London: College Publications.

[73] Priest, G. (2008), *Introduction to Non-Classical Logic: From If to Is*, Cambridge: Camnridge University Press.

[74] Priest, G. (2009), 'Conditionals: a Debate with Jackson', ch. 13 of I. Ravenscroft (ed.), *Minds, Worlds and Conditionals: Themes from the Philosophy of Frank Jackson*, Oxford: Oxford University Press.

[75] Priest, G. (2010a), 'Non-Transitive Identity', ch. 23 of R. Dietz and S. Moruzzi (eds.), *Cuts and Clouds: Vagueness, its Nature and its Logic*, Oxford: Oxford University Press.

[76] Priest, G. (2010b), 'Inclosures, Vagueness, and Self-Reference', *Notre Dame Journal of Formal Logic* 51: 69-84.

[77] Priest, G. (2010c), 'Hopes Fade for Saving Truth', *Philosophy* 85: 109-140.

[78] Priest, G. (2012a), 'The Sun may, Indeed, not Rise Tomorrow: a Reply to Beall', *Analysis* 72: 739-41.

[79] Priest, G. (2012b), 'Read on Bradwardine on the Liar', pp. 155-61 of C. D. Novaes and O. Hjortland (eds.), *Insolubles and Consequences: Essays in Honour of Stephen Read*, London: College Publications.

[80] Priest, G. (2014a), 'Contradictory Concepts', ch. 10 of E. Weber, D. Wouters, and J. Meheus (eds.), *Logic, Reasoning and Rationality*, Dordrecht: Springer.

[81] Priest, G. (2014b), 'Revising Logic', ch. 12 of P. Rush (ed.), *The Metaphysics of Logic*, Cambridge: Cambridge University Press.

[82] Priest, G (2014c), 'Plurivalent Logic', *Australasian Journal of Logic* 11, Article 1, http://ojs.victoria.ac.nz/ajl/article/view/1830.

[83] Priest, G. (2014d), *One*, Oxford: Oxford University Press.

[84] Priest, G. (2015), 'None of the Above: the Catuṣkoṭi in Indian Buddhist Logic', ch. 24 of J.-Y. Beziau, M. Chakraborty, and S. Dutta (eds.), *New Directions in Paraconsistent Logic*, Dordrecht: Springer.

[85] Priest, G. (2016a), 'Logical Disputes and the *a Priori*', *Logique et Analyse* 59: 347-66; and *Principios: Rivista de Filosofia* 23: 29-57.

[86] Priest, G. (2016b), 'Thinking the Impossible', *Philosophicl Studies* 173: 2649-62, and *Argumenta* 2: 181-94, 2017.

[87] Priest, G. (2017), 'What If: the Exploration of an Idea', *Australasian Journal of Logic* 14, https://ojs.victoria.ac.nz/ajl/article/view/4028/3574.

[88] Priest, G. (2018a), *The Fifth Corner of Four*, Oxford: Oxford University Press.

[89] Priest, G. (2018b), 'Some New Thoughts on Conditionals', *Topoi* 305: 369-77.

[90] Priest, G. (2018c), 'Paradoxical Propositions', *Philosophical Isues* 28: 300-7.

[91] Priest, G. (201+a), 'Dialetheic Metatheory', to appear.

[92] Priest, G. (201+b), 'A Note on Mathematical Pluralism and Logical Pluralism', to appear.

[93] Priest, G. (201+c), 'Kant's Excessive Tenderness for Things in the World, and Hegel's Dialetheism', to appear.

[94] Priest, G. (201+d), 'The Structure of Logical Dialectic', to appear.

[95] Priest, G. (201+e), 'Metaphysical Necessity: a Skeptical Perspective', *Synthese*, to appear.

[96] Priest, G. (201+f), 'Berry's Paradox... Again', to appear.

[97] Priest, G. (201+g), 'Some Contemporary Solutions to Some Medieval Problems about the Instant of Change', to appear.

[98] Priest, G. (201+h), 'Logical Theory Choice: the Case of Vacuous Counterfactuals', to appear.

[99] Priest, G. (201+i), 'Metaphysics and Logic: an Observation in Metametaphysics', to appear.

[100] Priest, G., and Routley, R. (1989a), 'Systems of Paraconsistent Logic', ch. 5 of Priest, Routley, and Norman (1989).

[101] Priest G., and Routley, R. (1989b), 'Applications of Paraconsistent Logic', ch. 13 of Priest, Routley, and Norman (1989).

[102] Priest, G., and Routley, R. (1989c), 'The Philosophical Signficance and Inevitability of Paraconsistency', ch. 18 of Priest, Routley, and Norman (1989).

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

[104] Priest, G., Siderits, M., and Tillemans, T. (2010), 'The Truth(s) about the Two Truths', ch. 8 of the Cowherds, *Moonshadows*, New York, NY: Oxford University Press.

[105] Priest, G., and Thomason, N. (2007), '60% Proof: Proof, Lakatos and Paraconsistency', *Australasian Journal of Logic* 5: 89-100.

[106] Pynko, A. (2010), 'Gentzen's Cut-Free Calculus versus the Logic of Paradox', *Bulletin of the Section of Logic* 39: 35-42.

[107] Read. S (2000), 'Harmony and Autonomy in Classical Logic', *Journal of Philosophical Logic* 29: 123-54.

[108] Read, S. (2008), 'The Truth Schema and the Liar', pp. 205-25 of S. Rahman, T. Tulenheimo, and E. Genot (eds.), *Unity, Truth and the Liar: the Modern Relevance of Medieval Solutions to the Liar Paradox*, Berlin: Springer.

[109] Restall, G. (1999), 'Negation in Relevant Logics (How I Stopped Worrying and Learned to Love the Routley Star)', pp. 53-76 of D. Gabbay and H. Wansing (eds.), *What is Negation?*, Dordrecht: Kluwer Academic Publishers.

[110] Restall, G. (2000), *Introduction to Substructural Logic*, London: Routledge.

[111] Ripley, D. (2001), 'Contradictions at the Borders', pp. 169-88 of R. Nouwen, R. van Rooij, U. Sauerland, and H. C. Schmitz (eds.), *Vagueness in Communication*, Berlin: Springer.

[112] Routley, R. (1977), 'Ultralogic as Universal', *Relevant Logic Newsletter* 2: 50–89, 138–75; reprinted as an appendix to *Exploring Meinong's Jungle and Beyond*, Canberra: Research School of Social Sciences, ANU.

[113] Routley, R. (1984), 'The American Plan Completed: Alternative Classical Style Semantics without Stars', *Studia Logica* 43: 131-58.

[114] Russell, B. (1908), 'Mathematical Logic as Based on the Theory of Types', *American Journal of Mathematics* 30: 222-62.

[115] Shapiro, S. (2002) 'Incompleteness and Inconsistency', *Mind* 111: 817-32.

[116] Shapiro, S. (2014), *Varieties of Logic*, Oxford: Oxford University Press.

[117] Shepard, R. (1964), 'Circularity in Judgments of Relative Pitch', *Journal of the Acoustical Society of America*, 36: 2346-53.

[118] Slater, B. H. (1995), 'Paraconsistent Logic?', *Journal of Philosophical Logic* 24: 451-4.

[119] Slater, B. H. (2007), 'Dialetheias are Mental Confusions', pp. 459-66 of J. Y. Béziau, W. Carnielli, and D. Gabbay (eds.), *Handbook of Paraconsistency*, London: College Publications.

[120] Spade, P. V., and Read, S. (2017), 'Insolubles', in E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/insolubles/.

[121] Tedder, A. (2015), 'Axioms for Finite Collapse Models of Arithmetic', *Review of Symbolic Logic* 8: 529-39.

[122] Tennant, N. (2015), 'A New Unified Account of Truth and Paradox', *Mind* 124: 571-605.

[123] Tenny, J. (1969), 'For Ann (Rising)', https://www.youtube.com/watch?v=bbKbE8y95sg.

[124] Urquhardt, A. (1984), ‘The Undecidability of Entailment and Relevant Implication’, *Journal of Symbolic Logic* 49: 1059-73.

[125] Vander Laan, D. (2004), ‘Counterpossibles and Similarity, ch. 20 of F. Jackson and G. Priest (eds.), *Lewisian Themes: the Philosophy of David K. Lewis*, Oxford: Oxford University Press.

[126] Voevodsky, V. (2012), ‘What if Current Foundations of Mathematics are Inconsistent?’, Lecture, Institute of Advanced Studies, Princeton University, https://www.youtube.com/watch?v=O45LaFsaqMA&t=2687s.

[127] Weber, Z. (2010), ‘Transfinite Numbers in Paraconsistent Set Theory’, *Review of Symbolic Logic* 3: 71-92.

[128] Weber, Z. (2012), ‘Transfinite Cardinals in Paraconsistent Set Theory’, *Review of Symbolic Logic* 5: 269-93.

[129] Weber, Z. (2016a), ‘What is an Inconsistent Truth Table?’, *Australasian Journal of Philosophy* 94: 533-48.

[130] Weber, Z. (2016b), ‘Paraconsistent Computation and Dialetheic Machines’, pp. 205-21 of H. Andreas and P. Verdee (eds.), *Logical Studies of Paraconsistent Reasoning in Science and Mathematics*, Dordrecht: Springer.

[131] Weber, Z., Badia, G., and Girard, P. (2016), ‘What is an Inconsistent Truth Table?’, *Australiasian Journal of Philosophy* 94: 533-48.

[132] Welch, P. (2008), ‘Ultimate Truth via Stable Truth’, *Review of Symbolic Logic* 1: 126-42.

[133] Welch, P. (2011), ‘Truth, Logical Validity, and Determinateness’, *Review of Symbolic Logic* 4: 348-59.

[134] Williamson, T. (2006), ‘Absolute Identity and Absolute Generality’, pp. 369-89 of A. Rayo and G. Uzquiano (eds.), *Absolute Generality*, Oxford: Oxford University Press.

[135] Williamson, T. (201+) ‘Counterpossibles in Metaphysics’, to appear in B. Armour-Garb and F. Kroon (eds.), *Philosophical Fictionalism*.

[136] Wittgenstein, L. (1968), *Philosophical Investigations*, Oxford: Basil Blackwell.

[137] Young, G. (2005), *Revenge: Dialetheism and its Expressive Limitations*, PhD Thesis, University of Glasgow.