Dialetheism, logical consequence and hierarchy

BRUNO WHITTLE

Dialetheism is defined by Graham Priest to be the view that there are true contradictions.[1] It is supposed to offer treatments of the semantic paradoxes that avoid the problems faced by more orthodox resolutions. The advantage of these treatments is supposed to be that they avoid the sort of appeal to a hierarchy of languages or concepts that more orthodox resolutions seem invariably to have to make. For since a dialetheist can simply accept as sound the derivations of contradictions involved in the paradoxes, there is no need for him to invoke a hierarchy to block these derivations.

In this article I argue that dialetheists have a problem with the concept of logical consequence. The upshot of this problem is that dialetheists must appeal to a hierarchy of concepts of logical consequence. Since this hierarchy is akin to those invoked by more orthodox resolutions of the semantic paradoxes, its emergence would appear to seriously undermine the dialetheic treatments of these paradoxes. And since these are central to the case for dialetheism, this would represent a significant blow to the position itself.

In §1 I explain why and how a dialetheist needs to be able to talk about logical consequence.[2] In §2 I argue that there are in fact severe restrictions upon how exactly a dialetheist can talk about logical consequence. These restrictions stem from a version of Curry's paradox. I then argue in §3 that a dialetheist must appeal to a hierarchy of concepts of logical consequence, and, further, that each of these concepts is dialetheically unobjectionable. The justification of this latter claim involves proving that the addition of these concepts together with natural rules for them conservatively extends dialetheic logic. This is proved in the appendix.

1. Talking about logical consequence

Like any other philosophical logician, a dialetheist wants to be able to talk about logical consequence. In particular, he wants to be able to express the fact that such-and-such follows logically from so-and-so. To this end, a dialetheist needs a 2-place connective that expresses such a

[1] See 2002: 4. For an extended defence of dialetheism see Priest 1987.

[2] By 'talking about logical consequence' I mean simply talking about whether such-and-such follows logically from so-and-so.

ANALYSIS 64.4, October 2004, pp. 318–26. © Bruno Whittle

concept of logical consequence.[3] Why does a dialetheist, or anyone else for that matter, need a connective that allows him to express such facts about logical consequence? Principally to be able to state instances of logical rules or, schematically or using variables, to be able to state the logical rules themselves. Indeed, unsurprisingly, Graham Priest talks frequently about what follows in dialetheic logic from what and what does not.

So let the dialetheist decide to use $\Rightarrow$ as this 2-place logical consequence connective. A non-technical English paraphrase of $\alpha \Rightarrow \beta$ would be something like: If α, then it follows logically that β. For these purposes I take logical consequence to include truth-theoretic consequence. For example, given that the dialetheic theory of truth includes the unrestricted truth schema, I take it that if τ is an individual constant naming α, then α is a dialetheic logical consequence of $T(\tau)$ and vice versa.[4] One could make the principal points that follow without using 'logical'. All that is essential is that this notion of consequence is a natural and important one.

As already noted, the main selling point of dialetheic treatments of the semantic paradoxes is their supposed avoidance of the problems faced by more orthodox resolutions as a result of their appeals to hierarchies. That is, their *raison d'être* is to eliminate any need for an object/meta-language distinction.[5] A dialetheist is thus going to want inferences whose validity depends on the meaning of $\Rightarrow$ to themselves be describable using $\Rightarrow$. So the notion of logical consequence captured by $\Rightarrow$ should include consequences which depend on the meaning of this very connective. In the next section I show that the dialetheist is not going to be able to give such an interpretation to $\Rightarrow$.

2. *Restrictions on talking about logical consequence*

Since the dialetheist is going to want to be able to describe using $\Rightarrow$ inferences that themselves depend on the meaning of $\Rightarrow$, he is going to

[3] The dialetheist could, of course, use a 2-place predicate instead, in which case he could take the predicate to apply to sets of sentences, instead of or as well as sentences. The arguments in this article would go through in much the same way if one considered such predicates rather than connectives. Indeed, if one used predicates one could at some points give even more direct arguments, since one would then be able to construct self-referential sentences using just the predicates in question, without having to use a truth predicate. I do not need the extra power that using such predicates would afford, and so I stick to the simpler case of connectives.

[4] For Priest on the truth schema see 1987: *passim*.

[5] For example, Graham Priest writes, '... the whole *point* of the dialetheic solution to the semantic paradoxes is to get rid of the distinction between object language and meta-language' (1990: 208).

want to accept every instance of (1) (I use α, β etc. as schematic letters for sentences).

(1) $(\alpha \wedge (\alpha \Rightarrow \beta)) \Rightarrow \beta$

After all, (2) is clearly a valid rule of inference if $\Rightarrow$ expresses a notion of logical consequence, dialetheic or otherwise.

(2) $\alpha \wedge (\alpha \Rightarrow \beta) / \therefore \beta$.

For (2) just records that β always follows from α together with the fact that if α then it follows logically that β. Of course, (2) is simply modus ponens for the connective $\Rightarrow$.[6]

However, no right-minded dialetheist can accept every instance of (1). To see why, let the individual constant c name the sentence $T(c) \Rightarrow \bot$ (for these purposes $\bot$ can be taken to stand in for any dialetheic absurdity, e.g. $0 = 1$).[7] (3) is then an instance of (1).

(3) $(T(c) \wedge (T(c) \Rightarrow \bot)) \Rightarrow \bot$

Substituting $T(c)$ for the sentence c names, i.e. $T(c) \Rightarrow \bot$, then gives (4).

(4) $(T(c) \wedge T(c)) \Rightarrow \bot$

I assume that a dialetheist will insist on being able to substitute $T(\tau)$ for α and vice versa in such a context, where τ is an individual constant that names the sentence α. Objecting to this move would seriously undermine the advantage the dialetheist perceives his position to have in virtue of its acceptance of the unrestricted truth schema. For, given such an objection, it would start to look as if, although dialetheism delivers every instance of the truth schema, the relevant biconditional must only express a very weak equivalence, given the failure of the substitution in question. So it

[6] It should be noted that in dialetheic logic modus ponens is invalid for the material conditional $\supset$, where $\alpha \supset \beta$ is defined as $\neg\alpha \vee \beta$ (see Priest 1987: 94–95). However, for this reason a dialetheist will reject the idea that $\supset$ is any sort of conditional at all: as Priest puts it, 'Any conditional worth its salt ... should satisfy the *modus ponens* principle' (1987: 103). Modus ponens would also be dialetheically invalid for a connective $\Rightarrow^*$ defined as follows: $\alpha \Rightarrow^* \beta$ iff it is logically true that $(\alpha \supset \beta)$. And similarly $\Rightarrow^*$ would be hopeless as a dialetheic logical consequence connective. For example, if γ is a standard Liar sentence, then for any sentence α one would have $\gamma \Rightarrow^* \alpha$. But the crucial point about dialetheism is that Liar sentences do not entail everything (see 1987: 103 and 6). Thus the fact that modus ponens is dialetheically invalid for $\supset$ and $\Rightarrow^*$ in no way indicates that a dialetheic logical consequence connective need not satisfy modus ponens.

[7] Of course, Graham Priest allows the construction of self-referential sentences by this sort of stipulation.

would begin to appear that the dialetheist could only hold on to the unrestricted schema by fundamentally weakening it.[8]

Substituting $T(c)$ for $(T(c) \wedge T(c))$ then gives (5).

(5) $T(c) \Rightarrow \bot$

Again I assume that the dialetheist will have no objection to this substitution. Obviously, if something is a logical consequence of $\alpha \wedge \alpha$ then it is a logical consequence of α, for classical or dialetheic logical consequence.[9]

But (5) is just the sentence c names, and so one can substitute $T(c)$ for (5) to get (6).

(6) $T(c)$

And from (5) and (6) one has (7) by $\wedge$-introduction.[10]

(7) $T(c) \wedge (T(c) \Rightarrow \bot)$

And then by rule (2) and (7) one has the absurdity (8).[11]

(8) $\bot$

Since all the moves taking one from (1) through to (8) appear to be dialetheically unassailable, the dialetheist would seem to be left with no choice but to reject (1).[12]

3. *A dialetheic hierarchy*

So (1) must be rejected. Yet there is a clear sense in which even from a dialetheic viewpoint something like (1) must be right. After all, there are no instances of rule (2) that are not valid. However, one cannot in general record the fact that an instance of (2) is valid using $\Rightarrow$, for that would just give an instance of the rejected (1). This means that one cannot in general describe using $\Rightarrow$ logical consequences that rely on the meaning of $\Rightarrow$ itself.

[8] Indeed, even in the context of Curry's paradox, where it would appear that for a dialetheist something desirable must be forsaken, Priest does not consider giving up such substitutivity of $T(\tau)$ for α and vice versa. See 1987: 104.

[9] Again, even in the context of Curry's paradox, Priest does not consider giving up on such substitutivity. See 1987: 104.

[10] Priest of course accepts $\wedge$-introduction.

[11] This argument is a version of Curry's paradox (see Curry 1942). It is in fact most closely related to the version of that paradox given in Meyer et al. 1979.

[12] Of course this does not mean that the dialetheist must reject every instance of the schema (1).

But one must surely be able to record these facts about instances of (2) somehow. So one must presumably introduce another 2-place connective, say $\Rightarrow_1$, in order to be able to state these facts, i.e. to essentially play the role being played by '/∴' in (2). And the natural way in which to do this is to keep $\Rightarrow$ simply for describing logical consequences that do not rely on the meaning of $\Rightarrow$, and then to use $\Rightarrow_1$ for describing the more inclusive class of logical consequences that results from adding those that rely on the meaning of $\Rightarrow$.

This understanding of $\Rightarrow_1$ would give (among others) the rule (9) and the schema (10).

(9) $\alpha \wedge (\alpha \Rightarrow_1 \beta)/\therefore \beta$.

(10) $(\alpha \wedge (\alpha \Rightarrow \beta)) \Rightarrow_1 \beta$

(9) is just (2) but for $\Rightarrow_1$ rather than $\Rightarrow$. And (10) records the fact that every instance of rule (2) is valid, which is statable with $\Rightarrow_1$ since $\Rightarrow_1$ can describe those consequences that rely on the meaning of $\Rightarrow$.

An instance of (10) is (3′).

(3′) $(T(c) \wedge (T(c) \Rightarrow \bot)) \Rightarrow_1 \bot$

And (3′) then leads to (5′) as (3) led to (5) in the previous section.

(5′) $T(c) \Rightarrow_1 \bot$

But there is no way of getting from (5′) to the sentence c names, i.e. $T(c) \Rightarrow \bot$, and hence no way of getting from (5′) to $\bot$. Thus it would appear that accepting (10) does not lead to the problems brought about by accepting (1).

Of course, for the same reason that (1) had to be rejected so does (11).

(11) $(\alpha \wedge (\alpha \Rightarrow_1 \beta)) \Rightarrow_1 \beta$

And thus $\Rightarrow_1$ cannot describe all of those consequences that rely on the meaning of $\Rightarrow_1$ itself. This then means that one needs yet another connective $\Rightarrow_2$, where $\Rightarrow_2$ stands to $\Rightarrow_1$ as $\Rightarrow_1$ stood to $\Rightarrow$. But then the same issue will re-emerge and one will need another connective $\Rightarrow_3$, and so on. Ultimately one will need as well as $\Rightarrow$ a connective $\Rightarrow_n$ for each natural number $n \geq 1$, and even beyond.

In the appendix I show that the addition of these new connectives together with natural rules for them conservatively extends dialetheic logic. More precisely, I show that given any set of sentences $\Phi \cup \{\alpha\}$ such that $\Rightarrow$, $\Rightarrow_1$ etc. do not feature in any sentence in $\Phi \cup \{\alpha\}$, or in any sentence named by a sentence in $\Phi \cup \{\alpha\}$, etc., then α is derivable from Φ in the extension of dialethic logic only if α is already derivable from Φ in standard dialetheic logic. Thus the introduction of the new connectives does not allow one to derive any conclusion not involving the new connectives that one could not already derive. Showing this is important since otherwise there would be no guarantee that the new connectives do not allow one to derive sentences in the original language that should not be regarded as true, such as for example $\perp$. However, given this result, it follows that at no point can the dialetheist reject any of these notions as unintelligible or incoherent.[13,14]

4. *Conclusion*

As noted in §1, the whole point of dialetheic treatments of the semantic paradoxes is their supposed avoidance of the sorts of hierarchies that are appealed to by more orthodox resolutions. However, I have shown that even if hierarchies are avoidable when talking about truth, they are not avoidable when talking about logical consequence. Thus, the supposed main advantage of these treatments would appear to be seriously undermined.

Appendix

In this appendix I show that the standard notion of dialetheic consequence is conservatively extended by the addition of natural rules for the connectives $\Rightarrow$, $\Rightarrow_1$, $\Rightarrow_2$ etc.[15] By the standard notion of dialetheic consequence I mean that which holds between a set of sentences Φ and a sentence α just in case α is derivable from Φ using just rules for the standard dialetheic logical constants (i.e. not including $\Rightarrow$, $\Rightarrow_1$ etc.), together with the following rule.

(TSub) $\alpha \vdash \alpha'$.

In (TSub) α' is either the result of replacing an occurrence in α of some sentence β by $T(\tau)$, where τ is an individual constant that names β, or vice

[13] Showing the introduction of a connective to be legitimate by proving that this introduction results in a conservative extension of the original consequence relation was first suggested in Belnap 1962. Unsurprisingly, Priest seems to accept that yielding a conservative extension is sufficient for intelligibility (see 1990: 204–5).

[14] That no such dialetheic response is available is what distinguishes the argument against dialetheism found in this article from arguments that attempt to use strengthened Liar paradoxes or versions of Curry's paradox to render dialetheism trivial. See for example Parsons 1990, Everett 1994 or Beall 2001 (Priest's replies to the first two of these are respectively 1995 and 1996).

[15] The rules that I consider are sound for the intended interpretations of $\Rightarrow$, $\Rightarrow_1$ etc. given in §3. I do not discuss completeness.

versa.[16,17] I take some fixed naming relation between individual constants and sentences to be given, and I assume that for every sentence there is an individual constant that names it. I use '$\vdash$' for this standard notion of dialetheic consequence.

Now let $\vdash_0$ be the extension of $\vdash$ that results from adding the following two rules (for sentences α and β). (This is the simplest and most natural way of extending $\vdash$ and is sufficient for my purposes; others are possible.)

(⇒I) If $\alpha \vdash \beta$, then $\vdash_0 \alpha \Rightarrow \beta$.
(⇒E) $\alpha \wedge (\alpha \Rightarrow \beta) \vdash_0 \beta$.

One wants (⇒I) simply because $\alpha \Rightarrow \beta$ is intended to mean that β is a logical consequence of α for a notion of logical consequence that depends on the meanings of the standard dialetheic logical constants, the truth predicate T and the names of sentences. (⇒E) is simply rule (2) of §2.

I will say that a first-order sentence α is *⇒-free* if α does not contain ⇒ and there is no sequence of individual constants $\tau_1, \dots, \tau_n$ ($1 \leq n \in \mathbf{N}$) such that (i) α contains τ_1; (ii) for each i with $1 \leq i < n$, τ_i names a sentence that contains τ_{i+1}; and (iii) τ_n names a sentence that contains ⇒.

One then has the following.

Theorem 0 Let $\Phi \cup \{\alpha\}$ be a set of ⇒-free sentences. If $\Phi \vdash_0 \alpha$ then $\Phi \vdash \alpha$.

Proof It is sufficient to show that, given any proof from Φ of a sentence β that terminates in an application of (⇒E), and that only uses (⇒E) once, one can construct a proof of β from Φ that does not use (⇒E) at all (β may here be ⇒-free or otherwise). So let x be such a proof of β, and let γ be the other sentence involved in the application of (⇒E) with which x terminates. x must contain a subproof of $\gamma \wedge (\gamma \Rightarrow \beta)$. And thus x must contain a subproof of $\gamma \Rightarrow \beta$ – or, by using $\wedge$-elimination, x can easily be modified so as to contain a subproof of $\gamma \Rightarrow \beta$, so I can assume that x contains such a subproof.

Either this subproof of $\gamma \Rightarrow \beta$ employs (⇒I) to infer $\gamma \Rightarrow \beta$ or $\gamma' \Rightarrow \beta'$ where β' and γ' are obtainable from β and γ (resp.) by repeated applications of (TSub), or this subproof does not so employ (⇒I).

[16] I use (TSub) rather than the T-schema simply to avoid having to enter into the vexed question of exactly which dialetheic biconditional the T-schema should be formulated with (see Priest 1987: 102–16). The T-schema results from (TSub) as long as one has $\alpha \leftrightarrow \alpha$ for any sentence α.

[17] A dialetheist should accept (TSub) for the reasons given in §2. And indeed Priest does (see 1987: 103–4).

First suppose that it does not so employ ($\Rightarrow$I). It follows that as far as the instances of rules used in this subproof are concerned, there is nothing to distinguish $\gamma \Rightarrow \beta$ from any other sentence (since Φ is $\Rightarrow$-free).[18] Thus it follows that this subproof can be modified to give a proof of any given sentence δ. But any such a proof of δ would use only standard dialetheic rules together with ($\Rightarrow$I). So in this case I am done by letting δ be β.

So now suppose that this subproof of $\gamma \Rightarrow \beta$ does employ ($\Rightarrow$I) to infer $\gamma \Rightarrow \beta$ or $\gamma' \Rightarrow \beta'$ where β' and γ' are obtainable from β and γ (resp.) by repeated applications of (TSub). If it employs ($\Rightarrow$I) to infer $\gamma \Rightarrow \beta$, then it follows that $\gamma \vdash \beta$. And given that this subproof contains (or can easily be modified so as to contain) a proof of γ that does not use ($\Rightarrow$E), it follows that x can be modified so as to prove β without using ($\Rightarrow$E). Similarly if the subproof employs ($\Rightarrow$I) to infer $\gamma' \Rightarrow \beta'$ where β' and γ' are obtainable from β and γ (resp.) by repeated applications of (TSub). ■

Thus the addition of ($\Rightarrow$I) and ($\Rightarrow$E) to the standard dialetheic consequence relation $\vdash$ results in a conservative extension.

One can then extend $\vdash_0$ to $\vdash_1$ by adding the following two rules (for sentences α and β).

($\Rightarrow_1$I) If $\alpha \vdash_0 \beta$, then $\vdash_1 \alpha \Rightarrow_1 \beta$.
($\Rightarrow_1$E) $\alpha \wedge (\alpha \Rightarrow_1 \beta) \vdash_1 \beta$.

($\Rightarrow_1$I) records the fact that $\alpha \Rightarrow_1 \beta$ is intended to mean that β is a logical consequence of α for the notion of logical consequence that is captured by $\vdash_0$. ($\Rightarrow_1$E) is essentially just rule (9) of §3. Schema (10) of §3 follows immediately from ($\Rightarrow$E) and ($\Rightarrow_1$I).

One can define relations $\vdash_n$ for each positive $n \in \mathbf{N}$ analogously. And if one defines '$\Rightarrow_n$-free' by exact analogy with the definition of '$\Rightarrow$-free' (i.e. with $\Rightarrow_n$ in place of $\Rightarrow$), then the following can be proved in just the same way as theorem 0 ($1 \leq n \in \mathbf{N}$).

Theorem *n* Let $\Phi \cup \{\alpha\}$ be a set of $\Rightarrow_n$-free sentences. If $\Phi \vdash_n \alpha$ then $\Phi \vdash_{n-1} \alpha$. ■

And one then clearly has the following ($1 \leq n \in \mathbf{N}$).

Corollary *n* Let $\Phi \cup \{\alpha\}$ be a set of sentences that are $\Rightarrow$-free and $\Rightarrow_i$-free for each i with $1 \leq i \leq n$. If $\Phi \vdash_n \alpha$ then $\Phi \vdash \alpha$. ■

So for each $n \in \mathbf{N}$, $\vdash_n$ conservatively extends $\vdash$. One can similarly define $\vdash_\lambda$, for any ordinal λ, and prove it to be a conservative extension of $\vdash$

[18] The assumption that every sentence is named by at least one individual constant is required here.

analogously. Thus, as promised, the addition of these new connectives together with natural rules for them conservatively extends the standard dialetheic consequence relation.[19]

Magdalen College
Oxford OX1 4AU, UK
bruno.whittle@magd.ox.ac.uk

References

Beall, Jc. 2001. A priestly recipe for explosive curry. *Logical Studies* 7: 1–15.

Belnap, N. D. 1962. Tonk, plonk and plink. *Analysis* 22: 130–34.

Curry, H. B. 1942. The inconsistency of certain formal logics. *Journal of Symbolic Logic* 7: 115–17.

Everett, A. 1994. Absorbing dialetheia? *Mind* 103: 413–19.

Meyer, R., R. Routley and J. M. Dunn. 1979. Curry's paradox. *Analysis* 39: 124–28.

Parsons, T. 1990. True contradictions. *Canadian Journal of Philosophy* 20: 335–54.

Priest, G. 1987. *In Contradiction: A Study of the Transconsistent*. Dordrecht: Martinus Nijhoff Publishers.

Priest, G. 1990. Boolean negation and all that. *Journal of Philosophical Logic* 19: 202–15.

Priest, G. 1995. Gaps and gluts: reply to Parsons. *Canadian Journal of Philosophy* 25: 57–66.

Priest, G. 1996. Everett's trilogy. *Mind* 105: 631–47.

Priest, G. 2002. *Beyond the Limits of Thought*. 2nd ed. Oxford: Clarendon Press.

[19] I would like to thank Timothy Williamson for discussions of earlier versions, and I am grateful to an anonymous referee for written comments. I would also like to thank the Analysis Trust for a studentship held while this article was written.