

Why and how to be a Dialetheist

(Draft, August 2007)

Dialetheism is the claim that some contradictions are true. For anyone trained in standard logic and raised in the belief that already in antiquity Aristotle settled once and for all that there is the Law of Non-Contradiction dialetheism sounds not just false, but bizarre.

On the other hand people contradict each other quite often and a couple of theories have turned out to be inconsistent. Nevertheless the people who held inconsistent beliefs have not (at the time of holding these beliefs) believed just anything, as the standard rule of *ex contradictione quodlibet* would have it. Thus paraconsistent logics (logics that invalidate *ex contradictione*.. and thus can tolerate even provable contradictions) have gained interest and lots of them are investigated and explored nowadays.

Dialetheism is *strong paraconsistency* in the sense that one cannot just tolerate some contradictions, but one should endorse some of them. This certainly needs argument.

In this paper the first part gives one of the main arguments why to be a dialetheist.

Ever since its arrival dialetheism (the thesis that there *are* true contradictions) has been met with the proverbial “incredulous stare”, not only because of the inconsistent ontology of Routley’s “noneism” (Routley 1979), but also with respect to the dialetheist’s claim that one can knowingly believe and assert contradictions. Priest in the paper introducing his “logic of paradox” **LP** (Priest 1979) admits that the thesis of dialetheism is a dialetheia itself, and seems to be content with this. In his book *In Contradiction* (Priest 1987) he argues that one can avoid dialetheism being a dialetheia itself if one is prepared to give up contraposition for the conditional in *Convention (T)*. Nevertheless he defends that one can believe and assert contradictions. Up to now (see some of the papers in (Priest/Beall/Armour-Garb 2004) or (Field 200x)) criticism of dialetheism has focused on the problems what the status of dialetheism itself is and how it may be possible to believe knowingly contradictions. In the second part of this paper it is argued that within dialetheism the resources are available to claim that dialetheism is true only (i.e. not false at the same time). Furthermore there may be occasions on which it is rational to believe and/or even assert contradictions, without thereby positioning oneself on a slippery slope towards an attitude of “anything goes”.

I. Why to be a Dialetheist

The main motivation for dialetheism is *universality* as a feature of language and cognition.

Universality means that we are aiming at – and supposedly capable of – a theory of, say, language *in general*, that is not just of this or that language or languages of this or that formal structure. And this theory is expressed in language, so that at least some language can be its own meta-language (with respect to all interesting properties of that language, semantics included). Universality means as well that we use fundamental concepts like *denotation* or *true* unrestrictedly.

A language can talk consistently about its own *syntax*. This presupposes expressive resources to name expressions/terms of the language and to represent syntactic properties. Once a language contains a basic system of arithmetic (the system **Q**) this is feasible. Within such a language L one can give a general structural description of what is a well-formed expression, what expressions are generalizations, and one can even define what counts as a derivation or a proof in that language. *Being provable in L* as a syntactic property of a formula can be defined within language L . Even if there was a hierarchy of language levels in L , given that the levels have some index there are formula that consistently talk about all such indices or talk about indices that are above the level such a formula is on. Syntactic hierarchies in contrast to semantic hierarchies are not strictly downwards, even in standard logic.

A language L is semantically closed iff L is able to talk about its own semantics. The meanings of the terms of L can be given within L then.

If a language is semantically closed it can not only talk about its own expressions (by suitable names or quotation marks), but it can also apply semantic properties to these terms, and even to the terms that express semantic properties:

- (1) Sentence number (1) is grammatical.
- (2) This sentence contains six meaningful words.
- (3) The third sentence displayed in this list is true.

Here some formula refers to itself by a description/name occurring in that formula. An infamous example is the Liar:

- (λ) λ is false.

The Liar is a fixed point for the predicate “() is false” (or “not-true”), saying “I am false”. Now consider (λ): If λ is true, then λ is false, because the general term in λ “() is false” should apply to the singular term “ λ ”. If λ is false, then λ is true, because λ is just saying that it is false. So we get:

- (4) $\text{True}(\lambda) \equiv \text{False}(\lambda)$

or given that we have a two-valued logic where “false” is just the opposite of “true”:

$$(5) \quad \text{True}(\lambda) \equiv \neg \text{True}(\lambda)$$

So λ is an *antinomy*: a sentence α where we have a proof for α and $\neg\alpha$. Usually (e.g. given standard propositional logic **PC**) this means we also have a proof of $\alpha \wedge \neg\alpha$.

As an example let's walk through the two proofs in case of the Liar:

[Proof of $\neg\lambda$]. If λ is true, then λ is false (by simple properties of being true). This is inconsistent, i.e. the assumption of λ being true leads to inconsistency, so by the rule of Negation-Introduction (\neg I) we get that λ cannot be true

[Proof of λ]. If (λ) is false, then λ is true, just saying that it is false. This is once again inconsistent, i.e. the assumption of λ being false leads to inconsistency so by (\neg I) we get that λ cannot be false.

[Proof of $\lambda \wedge \neg\lambda$]. Taking the two proofs together by Conjunction-Introduction (\wedge I) we get:

$$(6) \quad \lambda \wedge \neg\lambda$$

respectively

$$(7) \quad \text{True}(\lambda) \wedge \neg\text{True}(\lambda)$$

There are more harmless semantically self-referential sentences, like (2). The Liar is the basic case of bad semantic self-referentiality. It bears its badness on its sleeve. It is not hidden. Within a language that is semantically closed semantic self-reference may be hidden, however. Suppose today you only utter a single statement

$$(8) \quad \text{What the pope declares today is true.}$$

and *as a contingent matter of fact* he only says that day:

$$(9) \quad \text{Everything N.N. says today is false.}$$

The Liar is an antinomy. Dialetheism claims that it cannot be prevented, since a natural language *is* a semantically closed language. Since, furthermore, the antinomy can be *proved*, it has to be *true*.

So the dialetheist has to show three things:

- I. The contradictions can be *proven* in a *sound* non-standard logic, if we use a semantically closed language.
- II. We *have to* use a semantically closed language.
- III. There is *no satisfactory alternative* to accepting the antinomies (i.e. the attempts to prevent them either fail or have consequences worse than dialetheism).

I will briefly outline the arguments for these three claims (for the details see Bremer 2005).

(ad I)

To reason from the provability of an antinomy to its truth we need at least one correct/sound paraconsistent logical system. There are several. Besides that the reasoning about the Liar employs some properties of negation and Tarski's *Convention (T)*

$$(T) \quad \text{"p" is true (in L) if and only if p.}$$

ascribing truth to the name of a sentence/statement¹ is true just in case we can use that sentence truly, i.e. it is true. This convention is not tied to a specific theory of truth, since it may be taken either as a minimal condition, a statement of correspondence or a disquotational analysis. Even if one does not hold Convention (T) to be true there are other semantic antinomies using even less controversial concepts like “applying to” for a predicate or “denoting” for a singular term (for example Grelling’s Paradox using the fixed point of the predicate “does not apply to itself”).

True contradictions are said to be sentences such that α and $\neg\alpha$ are true. How is that compatible with our concept of *negation*? How can we justify that “ \neg ” behaving thus still expresses *negation*?

An intuitive acceptable concept of negation should support:

- (MN) (i) If α is false, then the negation of α is true.
- (ii) If α is true, then the negation of α is false.

These conditions seem to express the idea that the negation of a sentence α expresses the opposite of α , and a couple of paraconsistent logics (like **LP**) satisfy these conditions. If “false” and “not true” are the same, the two conditions just give us the ordinary truth table.

(ad II) in the dialetheist's agenda:

The natural languages we use seem to be semantically closed. We can express the antinomies in these languages. Natural languages have the resources of naming and corresponding self-reference. Semantic properties can be expressed in natural languages. *Prima facie* natural languages, therefore, are the very paradigm of universal languages. If somebody wants to deny this he has the burden of proof. Our ordinary conception of our language had to be seriously mistaken then! Usually the arguments *against* the semantic closure are based just on the antinomies.

Especially philosophy cannot restrict itself to non-universal languages. Philosophy does not want to deal only with the structure or conditions of talking in some specific language or languages of some kind, but aims at a theory of the basic structures and conditions of having a language *in general*. This requires the corresponding resources to express the universal claims. Universal theories of meaning, truth, knowledge etc. were not to have if we can talk only from some meta-language “down” to some distinct object-language. A general statement like

- (10) Knowledge is true belief.

would be not well-*formed*.

But these are the very theories that philosophy is after. And notwithstanding their lip-service to hierarchy solutions of the antinomies most philosophers propose their *general* theories of meaning, truth, belief, reference, knowledge etc. They are right to do the latter, since we have

¹ For the discussions here there is no important difference between statements, sentences and closed formula. So whereas strictly speaking only a sentence used in an *assertoric utterance* to make a *statement* is true or false I keep to the loose common standard of talking of the truth of sentences.

such universal concepts. We can investigate and formalize the logical structures of any natural languages. That is one of the central tenets of logic and formalization. We not only talk about properties of all (natural) languages, it seems even incoherent whether there could be two completely incommensurable languages. Our concept of language, therefore, involves unity and universality. There has to be a set of properties defining what a language is. These properties are preserved in change or translation. Without semantic closure we would not be *able* to elucidate a concept that we seem to have! So I take it that we need semantic closure. No-one, but dialetheism seems to be able to deliver it.

ad (III) of the dialetheist's agenda:

There are two major alternatives to paraconsistent treatments of the antinomies:

(IIIa) many-valued semantics or truth-value gaps

(IIIb) the hierarchy of semantic meta-languages

Both fail, but for different reasons.

(IIIa)-type solutions solve some antinomies, but the linguistic resources they employ in their formal framework are sufficient to generate new versions of the antinomies. (IIIb)-type solutions result in an outrageous pragmatic self-contradiction. Let us turn to (IIIa) first:

It has been claimed that the problems with the antinomies show that the crucial sentence has no truth value whatsoever or some further value besides “true” and “false”. Antinomies, however, are no problem tied to bivalence. If the Liar sentence λ is taken as neither true nor false the reasoning concerning sentence λ does not go through. That is right. The problem is that the *linguistic framework* employed to solve this antinomy is sufficiently rich to allow for new versions of the old antinomies, like the Liar. I take a linguistic framework to consist of both the language defined as well as its meta-linguistically expressed distinctions and semantics. Three-valuedness, for example, is part of some linguistic framework, given the interpretation of formulas in some semantics.

A general *hypothesis* of dialetheism is:

(LFT) A linguistic framework which is rich enough to avoid some of the antinomies, generates its own versions of them.

This hypothesis is rather vague, but it can be illustrated on several versions to deal with antinomies. With respect to three-valued approaches, one can introduce a *strengthened* Liar, (λ'). If the (old) Liar (λ) is neither true nor false, it is **not** true. A three-valued framework has two values (say “false” and “undetermined”) opposite to “true”. But now we can *say* that (λ) is on this opposite side, i.e. **not** true. This gives us a strengthened Liar:

(λ') λ' is not true.

We can argue again by cases – three cases now – and get a *new antinomy*:

(10) $\text{True}(\lambda') \equiv \text{Not-True}(\lambda')$

A framework involving three truth values involves the implicit or explicit validity of some principle like

(11) $\text{Not-True}(\alpha) \equiv \text{False}(\alpha) \vee \text{Undetermined}(\alpha)$

Even if "Not-True" is not *introduced* as a truth-value within the semantics, the semantic framework has *the expressive power* to introduce this notion. What we do here is to re-introduce a bifurcation within the realm of the truth-values. "Not-True" works like the notion of falsity in ordinary bivalent semantics. (This forced bifurcation is always available. The trick introducing the strengthened Liar is independent from the number of truth-values present.)

The strengthened Liar, therefore, might reside on the language level at which we can express "Not-True", and given the framework there has to be *some* such level.

Thus (IIIa)-type solutions won't work! What is wrong with the hierarchy solution [ad (IIIb)]? The main problem is not that at first sight we cannot find these levels in ordinary language. That is a problem, since the assumption of these levels would mean that there is some hidden syntactic structure with no corresponding surface structure, although it is decisive for truth! Such an analysis would be a major revision of our understanding of our language! Nonetheless there are even deeper problems in store:

The main problem is that if there were these levels and if the theory was true, the very statement of the theory and its given explanation (speaking *in general* about *all* language levels) would be impossible. The hierarchy conception says we are always talking from some level in the hierarchy, and *at the same time* makes a general statement to the effect that constructing the semantics of a language (level) we just go one level up.

Let's look at the details. Suppose we had a hierarchy of truth-predicates. Each of them has an index *i*. Ascribing truth would be like:

(12) α is true-at-level-*n* (in *L*).

Now if indexes are numbers, we are able to talk about them in a sufficiently rich language, able to talk about its own *syntax*. So we can built for each sentence α the expression "the-level-of- α ". Consider then the supposedly available sentence:

(13) (13) is not true-at-level-of-(13).

This sentence has to occur *somewhere* in the hierarchy on pains of the *inexpressibility* of some truths or talk about indices. Let us call this level "true-13"; then "true-at-level-of-(13)" is identical to "true-13". (T) applies at every level of the hierarchy for sentences of that level:

(14) True-13(α) if and only if α .

Applying (14) to (13) as a sentence at level 13 we get:

(15) True-13(13) if and only if (13).

And this yields a new antinomy:

(16) True-13(13) if and only if not true-13(13).

Being able to talk about the indexes simply re-introduces antinomies in the context of semantic vocabulary. The only solution to resist the argument above is to give up the assumption that (13) occurs somewhere in the hierarchy, has a level. Then, however, no

sentence that talks in this general fashion about indexes is possible. The following sentence would be impossible too:

(17) The truth predicate of a level n is defined at level $n+1$.

This sentence, however, can be expressed in natural languages, which are said to be captured. And (17) better be expressible if the theory of language levels is to be expressed at all! – If sentences like (17) were impossible the whole theory of (IIIb) would be *inexpressible itself!* Being unable to make statements about the hierarchy in general means that we are unable to understand the basic semantic concepts at all. So the hierarchy conception leaves us hanging in the air concerning our ability to understand semantic concepts at all! The hierarchy conception says we are always talking from some level in the hierarchy, and *at the same time* makes a general statement to the effect that constructing the semantics of a language (level) we just go one level up.

This is not just a contradiction. A contradiction one might suppose is what a dialetheist is to accept anyway. The situation is worse. If the theory is true – and, of course, as an adherent of the theory you believe it to be true – something is impossible *to do*, but you just do it! Given the theory the adherents are doing something impossible in the strict sense. What is done by the adherent of the hierarchy solution is completely mysterious. *Doing* the impossible seems to me even more bizarre than claiming that some contradictions are true. Even if you don't think so, the failure of the hierarchy solution is obvious: Dialetheism provides semantic closure – so it claims – at the cost of true contradictions; the hierarchy model implies either a mystery or a true contradiction *and* gets nothing for that, only the absence of semantic closure!

Dialetheism avoids mystery at the price of accepting some true contradictions. With respect to the general hypothesis concerning the failure of other solutions to the paradoxes one may refine the hypothesis by saying that either a linguistic framework will be rich enough to trade in new paradoxes for the old, or it will employ concepts *ineffable* within that framework or simply excludes linguistic resources obviously present in ordinary language. That is why one should be a dialetheist!

II. How to be a Dialetheist

Concerning a sentence α there are several levels of commitment. Consider:

- | | | |
|------|---|--|
| (18) | α is true. | [T [⌈] α ⌋, “T” being here a truth <i>predicate</i>] |
| (19) | Believing α . | [B α] |
| (20) | Rather concede α . | |
| (21) | Affirm α (assert that α is true). | [A α] |
| (22) | Abstain from an opinion on α . | [\neg B $\alpha \wedge \neg$ B $\neg\alpha$] |
| (23) | Disbelieve α . | [\neg B α] |
| (24) | Reject α . | [R α] |
| (25) | Believe the opposite of α . | [B $\neg\alpha$] |
| (26) | Assert the opposite of α . | [A $\neg\alpha$] |
| (27) | α is false. | [F [⌈] α ⌋, “F” being a falsity <i>predicate</i>] |
| (28) | The opposite of α is true. | [T [⌈] $\neg\alpha$ ⌋] |

Assertion as a speech act usually done in face of an audience commits one, at least *prima facie*, to provide reasons for one’s beliefs, if challenged to do so, whereas mere believe need not. There is a difference between abstaining from a judgement and disbelieving α if one seems to have reasons against believing α , but not against believing $\neg\alpha$. There is a difference between disbelief and rejection if disbelief is based on seeming to have reasons against believing α , and rejection on positively endorsing some reasons against α . If these reasons are taken as sufficiently strong, one believes $\neg\alpha$. “ \neg B α ” may cover disbelief, B $\neg\alpha$ then being “believing the opposite”. A $\neg\alpha$ is asserting the opposite. For this speech act the term “rejection” (R α) might be appropriate.²

Consistency principles then might be:

- | | |
|------|------------------------------------|
| (29) | $\neg(A\alpha \wedge A\neg\alpha)$ |
| (30) | $\neg A(\alpha \wedge \neg\alpha)$ |
| (31) | $\neg B(\alpha \wedge \neg\alpha)$ |
| (32) | $\neg(B\alpha \wedge B\neg\alpha)$ |

These principles, of course, seem to forbid anything like dialetheism. What thus seems intuitively so may not be fine grained enough, however, given the occurrence of true contradictions. And standard logic in its treatment of negation may level some distinctions that should be kept.

In standard logic (PC) rejection is equivalent to assertion of the opposite, since there is no 3rd value. Affirming α is rejecting $\neg\alpha$, and vice versa. Semantically we have:

² Henceforth, at least, “rejection” is used in that sense only.

α	$\neg\alpha$	$T\alpha$	$F\alpha$
1	0	1	0
0	1	0	1

α being not true means α is false, “false” being a synonym for the truth of a negation, expressed with “ \neg ”. The (T)-scheme is taken in its contrapositive form as well [$\neg\alpha \equiv \neg T\lceil\alpha\rceil$]. Therefore the standard logician endorses (29) – (32). Typically consistency of belief is demanded:

$$(33) \quad B\alpha \supset \neg B\neg\alpha \quad [\Leftrightarrow \neg B\alpha \vee \neg B\neg\alpha]$$

At least one of a sentence and its negation has to be disbelieved. Expressed with a truth predicate one demands

$$(34) \quad BT\lceil\alpha\rceil \supset \neg BT\lceil\neg\alpha\rceil$$

$$(35) \quad BT\lceil\alpha\rceil \supset \neg BF\lceil\alpha\rceil$$

For the standard epistemic logician

$$(36) \quad R\alpha \equiv AF\lceil\alpha\rceil$$

may be taken as the very definition of “rejection”.

Truth concerns what is the case whether we believe it or not. Belief concerns what we are willing to include in our inferring. What we believe we take into account in our reasoning (belief is cognitive).

Generally, being provided with reasons for α is seen as the basis for believing α , given that the reasons for some γ incompatible with α are not stronger.³

On a gullible approach to (perceptual) belief one believes every α one has no reasons against. The best backing for a belief α is a proof of α . Having reasons is superior to mere belief in the truth of α . Having no independent access to the (ultimate) truth of α going with reasons is the rational way, whatever the (ultimate) truth value of α is or turns out to be.

Given consistency and bivalence assumptions reasons against α may be reasons in favour of $\neg\alpha$, at least in non-empirical domains like semantics where a closed world assumption may be less idealized.

Typically it is taken to be rational to assent to [to affirm] what one believes. Assertion is to assent to or to affirm what one believes. If one has a belief α one also has the disposition to assert α . One does not need additional reasons to proceed from believing to asserting. On the other hand, asserting α is done by a speaker confronting an audience (assertion is pragmatic). Asserting α is done with a purpose in view of an audience, so that this purpose exceeds using α in one’s processes of deliberation (this being one’s self-satisfied belief that α). As an

³ This *proviso* depends on the consistency requirement not to have $B\alpha \wedge B\gamma$ with $\vdash(\alpha \supset \neg\gamma)$ [respectively: $\vdash(\neg(\alpha \wedge \gamma))$].

(speech) act with some purpose asserting α has to meet the basic conditions of successful action plans, like

- I. the purpose is not achieved anyhow *without* my action
- II. this specific action is *fit* to the purpose.

Asserting contradictions seems to fail both conditions, since there seems to be no specific commitment on the side of the speaker.

Given that standard logic runs into difficulties with antinomies also the principles supposedly governing belief, denial and asserting (the opposite) may need overhauling. Of special interest are now issues related to semantic closure and the formulation of the dialetheist position itself.

Conditions to be met by dialetheism are:

- I. Dialetheism as a thesis should be asserted as being *only/just true* (i.e. not being false at the same time).
- II. One should be able to say, without saying something false, that a true sentence/statement is true.
- III. One should be able to express the semantic properties of all sentences/statements (including the antinomies).

In meeting these conditions the dialetheist has to develop an understanding of denial and rejection which does not equate believing α with disbelieving $\neg\alpha$ and asserting α with rejecting $\neg\alpha$.

Reasons against α that are not reasons for $\neg\alpha$ may be reasons that undermine assumptions which usually support α . If α and $\neg\alpha$ are true at the same time reasons for $\neg\alpha$ are not – cannot be – reasons against α . (If neither α nor $\neg\alpha$ has to be true, reasons against α are not *per se* reasons in favour of $\neg\alpha$.) On the other hand, since following reasons is the rational way having reasons for α may lead one to accept α and having reasons for $\neg\alpha$ may lead one to accept $\neg\alpha$ at the same time.

Dialetheism claims that some contradictions are true. So we have some sentence λ with λ , $\neg\lambda$, $T\ulcorner\lambda\urcorner$, $T\ulcorner\neg\lambda\urcorner$, $F\ulcorner\lambda\urcorner$, $F\ulcorner\neg\lambda\urcorner$ to start with. The reasons for this are that these contradictions are provable given some unassailable principles and structures in a semantically closed language. Now, these antinomies being *true* and being *justified* as true, by proving them, give all the reasons to *believe* that they are true and thus to believe them (themselves). So a dialetheist should *believe*

(37) The Liar is true.

thus

(38) The Liar

and thus (by the definition of the Liar)

(39) The Liar is false.

Giving up believing what one has proven seems to be a desperate and *ad hoc* manoeuvre. So a dialetheist has inconsistent beliefs. She reasons *using both* $T\ulcorner\lambda\urcorner$ and $F\ulcorner\lambda\urcorner$ if necessary.

Paraconsistent logics can level the distinction between object and meta-language. A

semantically closed language not only is able to talk about its own expressions, but does contain at the same time its semantic expressions. These semantic expressions need not be taken as predicates (like a truth predicate applying to the quotation of a sentence), but can be taken as *operators* instead. One arrives at a paraconsistent language/logic which allows truth value talk without previously quoting the sentences which are evaluated.

To fulfil the condition of dialetheism being expressible we need bivalent truth operators working in the fashion of the following table:

α	$\neg\alpha$	$T\alpha$	$F\alpha$	$\Delta\alpha$	$\nabla\alpha$	$\circ\alpha$	$\bullet\alpha$
0	1	0	1	0	1	1	0
1	0	1	0	1	0	1	0
0,1	0,1	1	1	0	0	0	1

“ $\Delta\alpha$ ” says that α is true only, “ $\nabla\alpha$ ” that α is false only, “ $\circ\alpha$ ” says that α is consistent (i.e. has only one truth value), “ $\bullet\alpha$ ” says that α is contradictory. We can then say – and these being *just* true – that the Liar is true, false, not simply true, not consistent, and so on. “T” and “F” are *now* understood as *operators* applying to formulas/sentences not to quoted formulas/sentences.

Thus dialetheism can fulfil the traditional condition on any decent theory: that it claims to be just true (and not only as true as its negation). Dialetheism is thus no form of trivialism (that everything is true). The trivialist proposes $(\forall\alpha)(T\alpha \wedge T\neg\alpha)$ or $(\forall\alpha)(T\alpha \wedge F\alpha)$. The dialetheist claims $(\exists\alpha)(T\alpha \wedge F\alpha)$, but also $(\exists\alpha)(T\alpha \wedge \neg T\neg\alpha)$, and $(\exists\alpha)\nabla\alpha$. And given some formal system some formulas can be exhibited having these properties (e.g., defining a bottom particle \perp with $\nabla\perp$ being valid). \top can be defined as the top particle with $T(\alpha \vee \neg\alpha)$, being true only. The bottom particle \perp can be defined as $\nabla(\alpha \vee \neg\alpha)$, being false only. Note that – in contrast to even the intuitionist negation rules – $\perp \equiv (\alpha \wedge \neg\alpha)$ need not hold if α is a dialetheia, since then $T(\alpha \wedge \neg\alpha)$, and ∇ is incompatible with T.

To have and use the (T)-scheme at the same time as these operators (be it for the operator “T” or “ Δ ”) we need some revisions in the logic of the conditional, like giving up on the unrestricted validity of Contraposition. $T\alpha \wedge \nabla\alpha$ is a well-formed formula, but false only. The language of this version of dialetheism thus contains formula that can be evaluated only as being simply false. These formulas, of course, cannot be derived.

We do not need the details of all these restrictions here. The reader has only to know the general idea of paraconsistent logics and the idea of “adaptive logics” (Batens 1989, 2000) to restrict some rules to consistent sentences (respectively to retract some supposed consequences if the rules to derive them employed, against the restrictions, some inconsistent sentences). A paraconsistent logic like Priest’s **LP** can be developed into an adaptive logic with a restricted form of *Modus Ponens* and Contraposition (Priest 1991). Within

paraconsistent logics “logics of formal inconsistency” (Marcos 2005) employ consistency operators in the object language. Truth operators can then be added. Blending these approaches one can have an adaptive paraconsistent logic which combines the extensional and intuitive truth conditions of **LP** with the use of truth and consistency operators (Bremer 2005). We suppose here that the dialetheist uses some such logic. Adaptive logics employ standard logic in consistent context and with respect to consistent objects and use a paraconsistent logic for the inconsistent cases. They are adaptive in that one proceeds on the assumption that one deals with a consistent case only on explicit information that the context is inconsistent some supposed consequences have to be retracted. Practically this works by adding to natural deduction style derivation a further column in which one notes the consistency or normality assumptions or presuppositions that have to be made when employing some critical rules of inference. For example, the paraconsistent logic **LP** makes – as do paraconsistent logics typically – Disjunctive Syllogism invalid; since **LP**, further on, uses the standard material conditional this means that *Modus Ponens* is not valid in general; but it is valid on the assumption that the antecedent ϕ of the conditional $\phi \supset \psi$ used in an instance of *Modus Ponens* is a consistent statement. Thus noting the assumption $\circ\phi$ in the extra column of a derivation one can employ *Modus Ponens*, but once it turns out by the internal dynamics of drawing further consequences that ϕ was not consistent after all, the derived line and all lines dependent on it have to be retracted. We have to deal also with the failure of substitution of identicals for inconsistent objects. Identity elimination, ($=E$), has to be restricted to consistent objects. We define a consistency predicate “ $K()$ ” for objects (as a logical constant, of course) to do this:

$$(DK) \quad K(a) \stackrel{\text{def}}{=} \neg(\exists P)(P(a) \wedge \neg P(a))$$

Since we do not use a second order system here, we may employ (DK) in that way that we note $\neg K(a)$ in some line of a derivation if for the object named “a” we could have a line with an instance of the schema: $P(a) \wedge \neg P(a)$. Identity Elimination then takes the form:

$$\begin{array}{llll} n.<m> & P(a) & \dots & \Gamma \\ o.<k> & a = e & \dots & \Lambda \\ p.<m,k> & P(e) & (=E) n,o & \Gamma \cup \Lambda \cup \{K(e)\} \end{array}$$

where the column on the right takes down the sets of normality/consistency assumptions (or other presuppositions, cf. Bremer 2005: 224-36). The principal inconsistent object we are concerned with here is, of course, λ .

An example derivation looks like this:

$$\begin{array}{lll} 1.<1> & p & \text{Premise} \\ 2.<\diamond> & p \supset \neg \neg p & \text{PC} \\ 3.<1> & \neg \neg p & (\supset E)1,2 \quad \{\circ p\} \\ 4.<2> & \neg p \vee q & \text{Premise} \end{array}$$

5.<1,2>	q	(\vee E)3,4	{ $^{\circ}p, ^{\circ}\neg\neg p$ }
6.<2>	$p \supset q$	(\supset I)1,5	{ $^{\circ}p, ^{\circ}\neg\neg p$ }
7.< \diamond >	$(\neg p \vee q) \supset (p \supset q)$	(\supset I)2,6	{ $^{\circ}p, ^{\circ}\neg\neg p$ }

To return to the truth operators: Saying $T\lambda$ is thus simply true: $\Delta T\lambda$. This does not exclude that $F\lambda$ is also simply true: $\Delta F\lambda$.

Now it *seems* that saying of the Liar that the Liar is false is just what the Liar is saying

$$(40) \quad F\lambda \equiv \lambda$$

Then we might have

$$(41) \quad FF\lambda$$

and this contradicts $\Delta F\lambda$! But to derive (41) we use either

$$(42) \quad F\lambda = \lambda$$

taking the sentences as objects and expressing their identity, or

$$(43) \quad \vdash (F\lambda \equiv \lambda)$$

which may be a *petitio* in the argument under consideration, and then substitution of identicals or substitution of equivalents.

The equivalence thesis (43) may be wrong. And substitution of identicals is one of those inferences restricted to consistent objects (to which λ does not belong). Even if (43) is not wrong deriving $FF\lambda$ supposedly has to use some form of detachment, which again is restricted to consistent sentences (to which λ does not belong).

Let us take it that $F\lambda$ can be believed and – being bivalent – can be asserted. Asserting $T\lambda$ or $F\lambda$ certainly fulfils some purpose, be it in explaining dialetheism or in arguing with opponents of dialetheism.

What about λ itself? What could be the purpose of asserting λ when one could assert $\neg\lambda$ as well? Can asserting an antinomic sentence have any purpose at all?

Believing λ – as the dialetheist does – is not enough.

Given that the dialetheist is engaged in discussions about dialetheism it may be important to affirm her position by giving an example of what is a true contradiction. This can be done by affirming the antinomy itself, since we and the dialetheist take assertion to involve being convinced of the affirmed sentence being true (being *at least* true in the dialetheist's case). So if asserting α can be taken as asserting $T\alpha$ (not necessarily $\Delta\alpha$ in the dialetheist's case) and $T\lambda$ may be useful in a discussion about dialetheism, asserting λ has its place as well.

In memory of the distinction between object- and meta-language, dropped by the dialetheist, one may call this a *meta-assertion* of an antinomy. So there are occasions on which it is rational for a dialetheist to assert a contradiction.

Are there – apart from the just given purpose of uttering λ as a hidden/implied utterance of $T\lambda$ – other affirmative uses of λ ?

It seems not, since it seems difficult to come up with a purpose for affirming λ . Believing

both $T\lambda$ and $F\lambda$ (respectively λ and $\neg\lambda$) one may – it seems – as well use/affirm λ as $\neg\lambda$. But if any (non-meta-)usage of λ corresponds to a usage of $\neg\lambda$, there is no point in asserting λ , it seems. There seems to be nothing *specific* to be *said* by using λ ; even more so if a dialetheist accepts $\alpha \vee \neg\alpha$ as a tautology and rejects the use of disjunctive syllogism with antinomic sentences.

If there is no preference to affirm α in contrast to affirm $\neg\alpha$ why not affirm both? But again: Apart from conveying or displaying thus that α is taken as antinomic what is the supposed content of that assertion?

The semantic account of some predicates may speak of some quality/structure that entities have to which this predicate applies. Once *tertium non datur* is accepted – as it is by standard dialetheists – one either has to assume that $\neg\alpha$ contains the absence of the qualities/structures contained in α , which would make it difficult indeed to understand $\alpha \wedge \neg\alpha$ in a mildly realistic manner, or α and $\neg\alpha$ are seen as exclusive and exhaustive in the sense that they *both* contain some quality/structure the absence of both being (metaphysically) impossible.

Given a *substantial* theory of truth $T\alpha$ may convey some quality of α like corresponding to a fact, being rationally justified ... A substantial theory of falsity should accompany this theory, so that $\neg\alpha$ conveys some quality like the presence of a *negative fact* (!), being rationally refutable ... These qualities may co-occur! The dialetheist has to postulate some appropriate epistemological or metaphysical axioms then.

If the positive and the negative fact tied to the Liar are situated not in space and time but somewhere in our linguistic representation of the world, there may be room for a realist dialetheism which sees a purpose in asserting both λ and $\neg\lambda$. Given a metaphysics of this sort one can commit oneself by *one side* of a dialetheia. One takes up the commitment to argue that a corresponding structure is given. (This position has not to assume that the goal of affirmation is truth only, it is rather something being *at least* true.)

If the felicity conditions of affirmation/assertion entail that the purpose of affirmation is to claim something as being only true, then (by this alone) dialetheias are not affirmable. But why should one assume this?

The first reason seems to be that one is eager to exclude at the beginning a metaphysical picture of negative facts. The second reason, however, may rest on pragmatic felicity conditions of assertions as speech acts. Assertion requires to be pragmatically relevant that there is a commitment to something which has to exclude something else. If nothing is excluded by what I assert, I should not have bothered the effort. Now, in the typical presentation of antinomies (for example by arguing by cases $T\alpha/T\neg\alpha$) an antinomy α implies/entails $\neg\alpha$, and $\neg\alpha$ implies/entails α . Thus by either of them I assert what the other *says* as well. Therefore an account that bases the informational content of a sentence on what this sentence entails (cf. Priest 1987: 118) is of no help in these cases. So far this may point to the *arbitrariness* of which side of an antinomy is asserted only. By this reasoning one has no sufficient reason to affirm one side, and thus seems to be in some limbo of assertion.

One may think choosing just one side of the antinomy gets oneself out of this problem, an ontology of negative facts doing the rest. Asserting α (or $\neg\alpha$), however, has a point only if the facts corresponding to α and $\neg\alpha$, which by the mutual entailment of α and $\neg\alpha$ are put forward by either of them, are not exhaustive, it seems; something logically or semantically exhaustive being usually taken as having no informational impact because involving nothing to be excluded by it.

Negative facts have had a bad press in metaphysics.⁴ So one better had not oneself committed to them. Again, however, it seems that a *general* commitment to negative facts is as superfluous as a general acceptance of any old contradiction being true. The dialetheist accepts only very special “true contradictions”, namely those unavoidable given basic semantic or set theoretical concepts plus universality. The dialetheist, therefore, has to accept only very special negative facts. The failure of accounting for what the point of asserting a contradiction might be in terms of informational content or of what the two sides of the contradiction individually entail requires the more substantial account in terms of reference to distinct facts. In the case of Liar-like antinomies these facts consist in the negation of a sentence being as provable as the sentence itself. There is no further fact “behind” this. Since the proof is an existent *something* one may even speak of a positive fact here, like the intuitionist bases the claim for $\neg\alpha$ not on the absence of reasons for α , but on the (positive) proof of \perp from the premise α (cf. Priest 1987: 87).

Thus with respect to ordinary sentences (the truth of) $\neg\alpha$ may be the absence of (the truth of) α , but if α entails $\neg\alpha$ and vice versa, and both are of interest in as much as the fact corresponding to $\neg\alpha$ is not just the *absence* of the fact corresponding to α (as an “ordinary” supervenient negative fact would be) a substantial metaphysical assumption come to light: Both facts are substantial (and interesting), and it is a further substantial *metaphysical* fact that although they do not stand to each other like contradictory sentences do in **PC**, not both can be false only [corresponding to the theorem $\vdash(\neg(\nabla\alpha \wedge \nabla\neg\alpha))$].

Given that there is independent ground for $\neg\alpha$, accepting $\neg\alpha$ does not exclude accepting α . In contexts we know to be consistent we may reason to $\neg\alpha$ without independent grounds on the basis of $\neg T\alpha$ and *tertium non datur* (or some version of this disjunctive syllogism like reasoning). Since in case of antinomies accepting $\neg\alpha$ does not exclude accepting α , accepting $\neg\alpha$ should not be the same as rejecting α .

Rejecting α cannot be understood by a dialetheist as affirming $\nabla\alpha$. Rejecting α would thus be incompatible with affirming α (i.e. affirming $T\alpha$). One needs a distinction then between affirming $\nabla\alpha$ and affirming $F\alpha$ [$T\neg\alpha$]. Sticking with the usage employed above and –

⁴ At least negative *first-order* facts; the absence of all instances of a predicate understood as a supervenient negative fact has had a better press. Whether all negative facts corresponding to one side of a dialetheia are supervenient (i.e. non first-order facts) is not clear and may go against the spirit of, say, set-theoretical antinomies.

arguably – standard logic let us take affirming $F\alpha$ as *rejection* and affirming $\nabla\alpha$ as *denial* of α [$D\alpha$].

Whereas there are situations in which a dialetheist accepts both α and $\neg\alpha$, there are no situations in which a dialetheist accepts and denies α at the same time. Dialetheism does not accept just any contradiction. This is one reason – prejudices and puns to the side – why rational argument with a dialetheist is possible. As the foregoing distinction shows there is, furthermore, one kind of contradiction that (even) a dialetheist cannot support:

$$(44) \quad \neg(A\alpha \wedge D\alpha)$$

since $T\alpha$ and $\nabla\alpha$ are semantically incompatible.

Another simple point is that no-one (including the dialetheist) can have pragmatic contradictions: Speech acts being bodily movements that either occur or do not, there is no pragmatic parallel to having it both ways, i.e.

$$(45) \quad \neg(A\alpha \wedge \neg A\alpha)$$

This instance of the accepted tautology $\neg(\alpha \wedge \neg\alpha)$ expresses not only a semantic exclusion the dialetheist accepts (and sometimes nevertheless supersedes), but the absence of the mysterious feat of asserting something and not doing it at the same time. There is no pragmatic dialetheism (without a verbal manoeuvre of redefining “not asserting” on the lines of “asserting $\neg\alpha$ ”).

Having in mind these distinctions and the truth operators intuitively dialetheism allows for the truth (not necessarily the validity) of several sentences excluded by the consistency principles:

$$(46) \quad B\alpha \wedge B\neg\alpha \quad [\text{believing a contradiction}]$$

$$(47) \quad B(\alpha \wedge \neg\alpha)$$

$$(48) \quad BT\lambda \wedge BT\neg\lambda \quad [\text{being semantically explicit about } \lambda]$$

$$(49) \quad BF\alpha \wedge BF\neg\alpha$$

There is some spreading of believed or asserted contradictions in those paraconsistent logics in which $\neg(\alpha \wedge \neg\alpha)$ is a theorem or which are extended with the usual principle of closure of belief. Then we may have, for example:

$$(50) \quad B(\lambda \wedge \neg\lambda) \wedge B(\neg(\lambda \wedge \neg\lambda))$$

More controversial may be corresponding principles of assertion:

$$(51) \quad A\alpha \wedge A\neg\alpha$$

$$(52) \quad AT\alpha \wedge AT\neg\alpha$$

$$(53) \quad AF\alpha \wedge AF\neg\alpha$$

$$(54) \quad A(\alpha \wedge \neg\alpha)$$

$$(55) \quad A\alpha \wedge R\alpha \quad [\text{invalidating } A\alpha \supset \neg R\alpha \text{ and } R\alpha \supset \neg A\alpha]$$

Given the truth operators some new principles (and their duals), however, are in force now:

$$(56) \quad B\Delta\alpha \supset \neg BF\alpha$$

$$(57) \quad A\Delta\alpha \supset \neg AF\alpha$$

$$(58) \quad R\Delta\alpha \supset BF\alpha$$

(59) $\neg A(\Delta\alpha \wedge F\alpha)$

(60) $\neg B(\nabla\alpha \wedge T\alpha)$

Thus without sliding into mystery or being silenced one can be a dialetheist and claim some crucial antinomies to be true. Dialetheism itself is not a paradoxical statement, but the theory that fits the aspirations of an universally minded philosophy.

References

- Batens, Diderik (1989). "Dynamic Dialectical Logics", in: Priest, G./Routley, R./Norman, J. (Eds.) *Paraconsistent Logic*. Munich (Philosophia).
- (2000). "A Survey of Inconsistency-Adaptive Logics", in: Batens, D. et al. (Eds.) *Frontiers of Paraconsistent Logic*: Baldock (Research Study Press), pp. 49-73
- Bremer, Manuel (2005). *An Introduction to Paraconsistent Logics*. Bern et al. (Lang Publishers).
- Field, Hartry (200x). "Is the Liar Sentence Both True and False?", in: Beall, J./Armour-Garb, B. (Eds.) *Deflationism and Paradox*. Oxford (Oxford University Press), forthcoming.
- Marcos, Joao (2005). *Logics of Formal Inconsistency*. Campinas.
- Priest, Graham (1979). "The Logic of Paradox", *Journal of Philosophical Logic*, 8, pp. 219-41.
- (1987). *In Contradiction*. Dordrecht (Martinus Nijhof).
- (1991). "Minimal Inconsistent LP", *Studia Logica*, pp. 117-29.
- (1996). "Everett's Trilogy", *Mind*, 105, pp. 631-48.
- (2006). *Doubt Truth to Be a Liar*. Oxford (Clarendon Press).
- /Beall, J./Armour-Garb, B. (2004). (Eds.) *The Law of Non-Contradiction*. Oxford (Clarendon Press).
- Routley, Richard (1979). *Exploring Meinong's Jungle and Beyond*. Canberra.

Manuel Bremer, University of Düsseldorf, Germany