Logic and Logical Philosophy
Online First Articles (2024)
DOI: 10.12775/LLP.2024.008

Zach Weber

True, Untrue, Valid, Invalid, Provable, Unprovable

Abstract. There are many approaches to paraconsistency, ranging from the very moderate to the more radical. In this paper I explore and extend the more radical end of the spectrum, where there are truth-value gluts. In particular I will look at *paraconsistent metatheory* — the machinery of truth, validity, and proof — as developed in a glut-friendly paraconsistent setting. The aim is to evaluate the philosophical and technical tenability of such an approach. I will show that there are very significant technical challenges to face on this sort of radical approach, but that there is good philosophical support for facing these challenges.

Keywords: dialetheism; revenge paradoxes; paraconsistent metatheory

1. Introduction

1.1. From moderate to radical paraconsistency

There are many approaches to paraconsistency (see Beall and Restall, 2005, p. 80; Weber, 2022). Some are very moderate, aiming to *extend* standard classical theories with a non-explosive consequence relation in certain places, analogous to the way the transfinite extends but does not alter the finite natural numbers (Carnielli and Coniglio, 2016, p. x). Moderate approaches deny the principle of *ex falso quodlibet* (EFQ) because of possible inconsistency in our minds, or our language, or discussive contexts, or other worlds, but hold that consistency-based theorizing in science and mathematics is on the whole reliable and correct. And they assume that any theory needs a classical representation (e.g., a non-triviality 'consistency' proof, relative to some classical system) to be acceptable. Other approaches to paraconsistency are very radical, aiming

Special Issue on Paraconsistency. Guest Editors: D. Fazio, T. Jarmużek, J. Malinowski
Received February 24, 2023. *Accepted* September 22, 2023. *Published online* January 29, 2024

to *recast* or *revise* standard classical theories (e.g. Routley, 1977). They deny EFQ because the 'consistency hypothesis' has failed[1] and theories must be reconsidered and rehabilitated on a paraconsistent basis, so some classical results may be rejected, and new ones accepted perhaps even without classical representation.

In this paper I will explore and extend the more radical end of the spectrum, where there are *gluts* or *dialetheias*,[2] developing themes from (Weber, 2021). The aim is to evaluate the philosophical and technical tenability of a very radical approach. The strategy is to argue that, if there is good reason to use a glutty theory, then there is good reason to use a glutty *meta*theory too; the key issue is how to go about doing so.

We will begin with basic notions of truth and falsity, and how these relate to un-truth. A textbook way of allowing gluts, as in Priest's (2008), is to distinguish falsity from un-truth: a proposition may be false in the sense of having a true negation, without also failing to be true. This is often done via some notational markers like $+$ and $-$, where, e.g., $\vDash^- A$ is distinct from $\nvDash^+ A$. Indeed, this is in part what appears to make it possible for there to be non-trivial inconsistency at all. But I will argue that this approach is problematic, and outline an alternative in which falsity implies un-truth — so some (glutty) propositions are both true and un-true. Extending these notions to semantic validity, we consider ways in which an argument may be thought of as both valid and invalid, and here highlight the importance of an absurdity constant $\perp$ in the language. Finally, we consider operationalizing these notions in terms of derivability, where some conclusions may be thought of as both provable and unprovable.

The aim in presenting all of this together, in panoramic snapshot, is to appreciate the philosophical coherence — some might say single-mindedness — of such an approach, by explicitly identifying some of its guiding principles and how they are repeatedly applied. These are what in Section 1.2 I call the ORDINARINESS and NON-CLASSICALITY hypotheses, and the DIALETHEIC PARACONSISTENCY thesis. The aim is also to admit that the current state of the program is far from a

[1] The mainstream hypothesis "[...] that all that can be spoken of or described (non-trivially) is consistent" (Priest et al., 1989, p. 4); the idea that this hypothesis has *failed* is floated in (Routley and Meyer, 1976).

[2] Some authors have sought to distinguish glut theory from dialetheism (see, e.g., Ficara, 2021). This is worth discussing further but for now I will treat the terms as synonymous.

completed utopia and will require some dedicated and creative future researchers. I will conclude that there are very significant technical challenges to face on this sort of radical approach, but that there is good philosophical support for facing these challenges.

1.2. A question, and two hypotheses

Let's begin by facing a skeptical (but sincere) question for the radical approach. An amalgam of questions and comments I've heard over the years says:

> Okay, *maybe* thought or language can be inconsistent, but the world and its truths are obviously consistent. Bridges (mostly) do not fall down and the shoes on my feet are not also non-shoes. What would 'radically' paraconsistent things even be supposed to look like?

I will sketch a response, guided by a few key principles. One is the

Ordinary-ness Hypothesis. 'Paraconsistent' objects are ordinary objects, described in a paraconsistent language and theorized with a paraconsistent logic.

This tells us not to look for some exotic or wild sorts of 'paraconsistent' entities to populate our theories. Rather, our theories are about everyday entities, like bridges and shoes; the theory just happens to be governed by a background paraconsistent logic. On this approach, paraconsistency is not about some "realm beyond the consistent" (Priest, 2006, p. 209). Maybe your shoes are so old and full of holes they really *are* both shoes and not shoes anymore (Beall and Colyvan, 2001, cf.[]). ORDINARINESS tells us to expect the expected.

The second hypothesis is the

'Non-classicality is not Classical' Hypothesis. Conventional standard theories are relative to classical logic. When conventional concepts, definitions, proofs and theorems are expressed in a paraconsistent language and established using only paraconsistently valid arguments, they may diverge from the classical case.

This tells us not to look for copies or replicas of classical theories, somehow delivered by non-classical logics. Such a goal, the so-called 'classical recapture', has long been pursued by paraconsistent researchers but as we will see in this paper, there are significant reasons to set new goals instead (cf. Sylvan and Copeland, 2000; Meadows and Weber, 2016). Why pursue a non-classical path just to regain what the classical already had? NON-CLASSICALITY tells us to expect the unexpected.

Putting these together, ORDINARINESS says to expect the expected with respect to experience, while NON-CLASSICALITY says to expect the unexpected with respect to classical theories. Between these two poles, a radical approach can still find balance. We will see that, e.g., a 'paraconsistent chair' is just an unsurprising chair — and *that* in itself is surprising.

1.3. Gluts

According to truth-glut theory (Beall, 2009) or *dialetheism* (Priest et al., 1983) some propositions are both true and false: there is some true proposition p such that $\neg p$ is true, too. Such p is true simpliciter, not (only) true in some model, though it is that, too. Such p is true in the actual world, whatever 'true in the actual world' happens to mean.

A reason for believing in gluts is that some of them appear at the end of what look like deductive proofs, and accepting that this is so provides direct resolution of famous paradoxes in logico-mathematics, especially the interrelated self-referential fixed-point phenomena that underly the famous limitive theorems; this is argued directly in (Priest, 2006, ch. 1, 2, 3) and indirectly (but to my mind even more compellingly) in (Priest, 2002a); see (Weber, 2021, ch. 0, 1). A reason this solution is appealing is that it seems to avoid the 'revenge' (all?) other solutions face. All approaches to the paradoxes seem eventually to contradict themselves, or else go terminally silent (which itself may be some kind of performative contadiction). Dialetheism contradicts itself, too — it explicitly endorses contradictions, including in the statement 'no sentence is both true and false' (see Section 1.5 below) — but it is the only approach where that is not a cost.[3]

In this sense dialetheism is not a 'solution' to the paradoxes — it does not make them 'go away' — but rather a response that attempts to accept the paradoxes and learn from them. If there is a fixed point knot at the bottom of semantics, set theory, recursion theory, and perhaps even arithmetic, as the years from Russell to Gödel and Turing seem to show, then the way forward, according to dialetheism, is to take this knot as a part of logico-mathematics to be studied and gripped, not untied.

---

[3] As Shapiro puts it, "An important advantage — perhaps the major advantage — of the dialetheic program is the possibility of a single uniform semantics [. . .]. [W]e do not need to keep running through richer and richer metalanguages in order to chase our semantic tails. [. . .] We embrace some contradictions in the semantics, and get it all from the start" (Shairo, 2002, p. 818), adding: "Or so says Priest".

1.4. Methodology: Paraconsistency and Meta-Paraconsistency

For dialetheism to be rationally coherent, there needs to be an accompanying paraconsistent logical system in which inconsistency does not spread to all propositions. This is encapsulated in the following truism.

Paraconsistency Thesis. Not everything is true. Not everything is true *even if* some contradictions are.

Combining with gluts, we get

Dialetheic Paraconsistency. Some but not all contradictions are true.

Following through on dialetheic paraconsistency as a program requires setting up notions of truth, satisfaction, validity, and provability that deliver results without *over*-delivering. To escape revenge and maintain its self-styled status as the only approach that does escape revenge, dialetheic paraconsistency must set up its own framework in a way that does not appeal to unavailable notions or otherwise run afoul of its own proscriptions. If a position eschews classical logic, but must use classical logic essentially in order even to state the position, that is prototypical failure.[4]

Let us at least try to do better. I would suggest that, since the modern logic project launched at the turn of the 20th century to rousing Hilbertian aspirations, we have grown accustomed to bad news and disappointment. All the big-name results are negative: first order languages cannot pin down intended models (Löwenheim-Skolem); axiomatic systems cannot reach the full set of truths (Gödel); truth is not even expressible (Tarski); many important problems are not computable, including the question of which problems are computable (Church, Turing); and so on. A thoroughly paraconsistent approach keeps alive the possibility of doing better, but only through genuine NON-CLASSICALITY.

How then to arrange an internally coherent dialetheic paraconsistent system? Perhaps unsurprisingly, we will focus on the role of *negation*, in the notions of truth, validity, and proof. To do this, we first need to fix the background framework for these developments, a logic and some mathematics to build logic and mathematics.

[4] A referee describes this as *logical constancy* — avoiding the 'inconstancy' or *incoherence* (in Sylvan's phrase) of a mismatch between philosophical motives and methodological approach. See (Weber, 2021, ch. 3).

1.5. Logic for the metatheory

There are many available paraconsistent logics. Here is an axiomatic presentation of a a Hilbert System for the logic BCK (cf. Ono and Komori, 1985, p. 173) *plus* a de Morgan negation (Badia et al., 2019), quantifiers, and identity.

The language is of first order logic with identity, containing logical operators $\&, \vee, \neg, \bot, \Rightarrow$ (where $\Rightarrow$ is not definable in terms of the other operators), quantifiers $\forall, \exists$, a binary relation $=$ and punctuation, all with usual formation rules.

1. $A \Rightarrow (B \Rightarrow A)$
2. $(A \Rightarrow (B \Rightarrow C)) \Rightarrow (B \Rightarrow (A \Rightarrow C))$
3. $\bot \Rightarrow A$
4. $(A \Rightarrow B) \Rightarrow ((C \Rightarrow A) \Rightarrow (C \Rightarrow B))$
5. $(A \Rightarrow C) \Rightarrow ((B \Rightarrow C) \Rightarrow (A \vee B \Rightarrow C))$
6. $(A \Rightarrow (B \Rightarrow C)) \Rightarrow (A\&B \Rightarrow C)$
7. $A \Rightarrow (B \Rightarrow A\&B)$
8. $A \Rightarrow A \vee B$
9. $B \Rightarrow A \vee B$
10. $\neg\neg A \Leftrightarrow A$
11. $\neg(A\&B) \Leftrightarrow \neg A \vee \neg B$
12. $\neg(A \vee B) \Leftrightarrow \neg A\&\neg B$
13. $A \vee \neg A$
14. $\forall x A(x) \Rightarrow Ay$
15. $\forall x(A \Rightarrow Bx) \Rightarrow (A \Rightarrow \forall x Bx)$
16. $\forall x(A \vee B) \Rightarrow (A \vee \forall x Bx)$
17. $Ay \Rightarrow \exists x A(x)$
18. $\forall x(A \Rightarrow B) \Rightarrow (\exists x A(x) \Rightarrow B)$
19. $\neg\exists x A \Leftrightarrow \forall x \neg A$
20. $\neg\forall x A \Leftrightarrow \exists x \neg A$
21. $\forall x(x = x)$
22. $\forall x \forall y(x = y \Rightarrow (Ax \Rightarrow Ay))$

Rule 1 If A and $A \Rightarrow B$ then B
Rule 2 If A then $\forall x A$

Then an argument from $A_0, \ldots, A_n$ to A is *valid*,

$$A_0, \ldots, A_n \vdash A$$

iff A is an axiom, or follows from axioms or $A_0, \ldots, A_n$ by rules.

The conjunction connective & is an intensional, multiplicative operator that residuates the conditional,

FACT 1. $A \,\&\, B \Rightarrow C \quad \textit{iff} \quad A \Rightarrow (B \Rightarrow C)$

(unlike the truth-functional lattice conjunction $\wedge$ discussed in Section 2.1 below). See (Weber, 2021, ch. 4) for further discussion. Two further important facts to note are that this logic obeys a *deduction theorem*,

FACT 2. $A_0, \ldots, A_n \vdash B$ *iff* $\vdash A_0 \,\&\ldots\&\, A_n \Rightarrow B$.

and that it is 'Curry paraconsistent',

FACT 3. $A \Rightarrow (A \Rightarrow B) \nvdash A \Rightarrow B$, *and* $A \nvdash A \,\&\, A$.

Versions of *structural contraction* are not valid, in the sense that one may have $A \,\&\, A \vdash B$ without $A \vdash B$, due to Curry's paradox (cf. Section 5.1).

The appeal of such a system is to extend the language with a two place relation $\in$ and a variable binding term forming operator $\{\cdot : \cdot\}$, for a *naive set theory*:

AXIOM 1 (Comprehension).
$\forall x(x \in \{y : A\} \Leftrightarrow A), \quad \forall x(x \notin \{y : A\} \Leftrightarrow \neg A)$

The negative form of comprehension is added manually (where '$x \notin y$' abbreviates $\neg(x \in y)$) since $\Rightarrow$ does not contrapose.[5]

The system *without* de Morgan negation (so deleting logical axioms 10-13) plus positive naive comprehension, which seems to be initially due to Grišin, has been proven non-trivial, due to the absence of contraction (cf. Petersen, 2000). Whether the addition of de Morgan negation is non-trivial or not is an open (classical) question (see Weber, 2021, ch. 3, §3) for discussion. This weaker system alone can be used to define and prove properties of the most basic widgets, such as

- subsets $x \subseteq y := \forall z(z \in x \Rightarrow z \in y)$
- singletons $x \in \{a\} \Leftrightarrow x = a, \quad x \notin \{a\} \Leftrightarrow x \neq a$
- ordered pairs $\langle a, b\rangle := \{\{a\}, \{a, b\}\}$ from which one proves the law

$$\langle a, b\rangle = \langle c, d\rangle \Leftrightarrow a = c \,\&\, b = d$$

[5] Given comprehension, $\perp$ in our language is definable as $\forall x \forall y x \in y$. And with this, transitivity (logic axiom 5) does give $(A \Rightarrow B) \Rightarrow ((B \Rightarrow \perp) \Rightarrow (A \Rightarrow \perp))$. But $\neg A$ is not definable as $A \Rightarrow \perp$. In particular, the latter does not follow from the former, because inconsistency does not necessarily explode — that is paraconsistency! For the role of $\perp$ in nearby languages (see Badia, 2016).

- products $X \times Y = \{\langle x, y\rangle : x \in X \ \&\ y \in Y\}$, and relations $Z \subseteq X \times Y$
- an indexing set $\{0, 1, 2, \ldots\}$ obeying some limited form of induction

The development of this machinery can be found in (Cantini, 2003; Terui, 2004). We can take these as further postulates for now, in the spirit of inspiring pioneer exercises like (Arruda and Batens, 1982).

Ideally, one would also want to add to this system an axiom of *extensionality*, asserting that $\forall z(z \in x \Leftrightarrow z \in y) \Leftrightarrow x = y$. However it is known that this, in terms of the BCK arrow, will trivialise (Weber, 2021, ch. 4). Without extensionality, we really have a *property* theory rather than a set theory. This can be dealt with by adding another conditional to the system, e.g., an implication connective from weak relevant logic as in (Brady, 2006), but this is not without problems (Field, 2020; Weber, 2020), and trying to solve them now would take us too far off track. With this flagged, for our purposes here we will only need to know when two sets are *not* identical — when they differ with respect to membership — and this can be stated as

AXIOM 2. $\forall x \forall y(x \neq y \Leftrightarrow \exists z(z \in x \ \&\ z \notin y) \vee \exists z(z \notin x \ \&\ z \in y))$

Remark 1. If the Russell set is both a member of itself and not, then it is not self identical either (even though it also is self-identical, by logical axioms).

There are many approaches to paraconsistent set theory: for LP set theory, see (Restall, 1992; Martinez, 2021); for non-transitive approaches, see (Ripley, 2015; Istre, 2017); for others, see (Libert, 2003), (Carnielli and Coniglio, 2016, ch. 8), (Batens, 2020). So this is one example, which seems to have some promise from field testing, but it is not presented here as final or definitive. It seems likely that there are improvements and additions (for example, restricted quantifiers (Beall et al., 2006; Badia et al., 2022), using sparingly below) that will in future make for a fuller account for framing (paraconsistent) mathematics.

2. True and Untrue

The goal is to use the meta-system just sketched to build up some standard internal machinery for logic, starting with the notion of truth in a model. First we will look at a a 'standard' way of presenting a glut-friendly paraconsistent system, as found in, e.g., (Priest, 2008) or (Beall and Logan, 2017), which is to distinguish between *falsity* and *untruth*. E.g. in presenting the relational semantics for FDE (first degree entailment), we find

> Note that it is now very important to distinguish between being false in an interpretation and not being true in it. (There is, of course, no difference in the classical case.) The fact that a formula is false (relates to 0) does not mean that it is untrue (it may also relate to 1). And the fact that it is untrue (does not relate to 1) does not mean that it is false (it may not relate to 0 either). (Priest, 2008, p. 143)

The whole paraconsistent enterprise, it might seem, is made possible by separating falsity and untruth, so that a sentence might be both true and false, but never both true and untrue. This distinction underwrites a *consistent* presentation of possible inconsistency.

I will set out how this is done, how it might be problematic, and how it might be undone. To keep things contained, we will make the 'object' level logic of study the well-known *logic of paradox* LP.

2.1. Models

Let us look at a standard presentation of first order LP with identity. Let us emphasize that, following standard procedure, what we are about to look at is built in *classical* logic and set theory. This is for display and comparison only.

For the language, we have variables x, y, z, constants a, b, c, $\ldots$, predicate symbols F^n, G^n, H^n, $\ldots$ for predicates of any arity $n > 0$, a special two-place identity relation $=$, connectives $\wedge, \vee, \neg$, quantifiers $\forall, \exists$, and brackets. (Note that the conjunction $\wedge$ here is not the & from above.) Variables and constants are terms t_0, t_1, $\ldots$. If t_0, $\ldots$, t_n are terms and F is an n-place predicate then $F(t_0, \ldots, t_n)$ is an atomic formula. If A, B are formulas and x is a variable, then $\neg A$, $A \wedge B$, $A \vee B$, $\forall x A(x)$ and $\exists x F(x)$ are formulas.

DEFINITION 1 (Standard Model Structures). A *(standard) model structure* $\mathcal{M} = \langle D, \mathscr{I} \rangle$ consists of a non-empty domain D and interpreation $\mathscr{I}$ such that

- If c is a constant, $\mathscr{I}(c) \in D$.
- If F is an n-place predicate, then

$$\mathscr{I}^+(F), \mathscr{I}^-(F) \subseteq \{\langle d_1, \ldots, d_n \rangle : d_1, \ldots, d_n \in D\}$$

such that the interpretation is exhaustive, $\mathscr{I}^+(F) \cup \mathscr{I}^-(F) = D$, but *without* insisting that it is exclusive (which would be written $\mathscr{I}^+(F) \cap \mathscr{I}^-(F) = \emptyset$).
- For identity, $\mathscr{I}^+(=) = \{\langle x, x\rangle : x \in D\}$, but without insisting $\mathscr{I}^-(=) = \emptyset$.

The interpretation of F consists of separate specifications of the extension and anti-extension of F, its members and non-members. It is assumed that these are exhaustive but not exclusive. Similarly it is not assumed that the anti-extension of $=$ is empty.

DEFINITION 2. For each interpretation, a *standard valuation* v relates formulas to truth values $\{\mathsf{t}, \mathsf{f}\}$, via base conditions

$$\begin{array}{lcl} \mathsf{t} \in v(F(t_0, \ldots, t_n)) & \text{iff} & \langle \mathscr{I}(t_0), \ldots, \mathscr{I}(t_n)\rangle \in \mathscr{I}^+(F) \\ \mathsf{f} \in v(F(t_0, \ldots, t_n)) & \text{iff} & \langle \mathscr{I}(t_0), \ldots, \mathscr{I}(t_n)\rangle \in \mathscr{I}^-(F) \end{array}$$

and by recursion,

$$\begin{array}{rcl} \mathsf{t} \in v(\neg A) & \text{iff} & \mathsf{f} \in v(A) \\ \mathsf{f} \in v(\neg A) & \text{iff} & \mathsf{t} \in v(A) \\ \mathsf{t} \in v(A \wedge B) & \text{iff} & \mathsf{t} \in v(A) \text{ and } \mathsf{t} \in v(B) \\ \mathsf{f} \in v(A \wedge B) & \text{iff} & \mathsf{f} \in v(A) \text{ or } \mathsf{f} \in v(B) \\ \mathsf{t} \in v(A \vee B) & \text{iff} & \mathsf{t} \in v(A) \text{ or } \mathsf{t} \in v(B) \\ \mathsf{f} \in v(A \vee B) & \text{iff} & \mathsf{f} \in v(A) \text{ and } \mathsf{f} \in v(B) \\ \mathsf{t} \in v(\forall x A) & \text{iff} & \mathsf{t} \in v(A(x/d) \text{ for all } d \in D \\ \mathsf{f} \in v(\forall x A) & \text{iff} & \mathsf{f} \in v(A(x/d) \text{ for some } d \in D \\ \mathsf{t} \in v(\exists x A) & \text{iff} & \mathsf{t} \in v(A(x/d) \text{ for some } d \in D \\ \mathsf{f} \in v(\exists x A) & \text{iff} & \mathsf{f} \in v(A(x/d) \text{ for all } d \in D \end{array}$$

This would seem to have the virtue of 'relaxing' the classical conditions (which can simply be re-obtained by replacing $\in$ with $=$ above). A valid argument is truth preserving, semantically, iff when all the premises are 'at least' true, then the conclusion is 'at least' true. This leaves room for gluts to be modeled consistently.

Remark 2. If the membership relation $\in$ and $\notin$ are from *classical* set theory then the above semantics are equivalent to a three valued functional presentation:

$$v(A) = \begin{cases} \{\mathsf{t}\} & \text{if } \mathsf{t} \in v(A), \mathsf{f} \notin v(A) \\ \{\mathsf{f}\} & \text{if } \mathsf{t} \notin v(A), \mathsf{f} \in v(A) \\ \{\mathsf{t}, \mathsf{f}\} & \text{if } \mathsf{t}, \mathsf{f} \in v(A) \end{cases}$$

If the background negation and membership relation are consistent, then this rules out anything being both true and untrue, e.g., $\mathsf{t} \in v(A)$ and $\mathsf{t} \notin v(A)$. And if that is to be ruled out, then at least one of the following needs to be blocked:

Exhaustion. If $\mathsf{t} \notin v(A)$ then $\mathsf{t} \in v(\neg A)$

Exclusion. If $\mathsf{t} \in v(\neg A)$ then $\mathsf{t} \notin v(A)$

Priest drops exclusion, as it "multiplies contradictions beyond necessity" (Priest, 2006, p. 71) since otherwise, if A is both true and false, then it would be true and untrue.

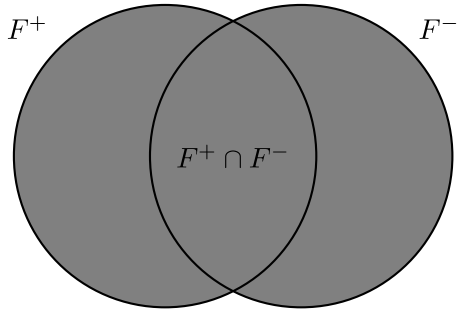
2.2. From classical to non-classical

The standard approach is to provide a *consistent* account of what non-trivial inconsistency might look like; but upon reflection, for a committed dialetheist it can start to seem self-defeating and revenge-inviting to attempt to maintain consistency at such a 'meta'-level. Does the standard approach restore coherence to gluts by removing the gluttiness of gluts? If truth and falsity do not contradict then why is a true/false pair a (true) contradiction?

The standard way opens paths to revenge, via strengthened liars formulated with classical negation (Omori and Weber, 2019). Assuming the meta-level is expressive enough, one can forsee a revenge sentence on the horizon, along the lines of, e.g., a property $\mathfrak{F}$ such that $\mathfrak{F}(x)$ iff $x \notin \mathscr{I}^+(\mathfrak{F})$. It makes a good deal more sense to take $\notin$ to be not classical. If paraconsistent dialetheism tells us to forswear classical negation to resolve the liar, then a paraconsistent dialetheist cannot revert back to classical negation to describe their own semantics! That would seem to violate the spirit of the project and undercut what it is for. If truth and untruth are absolutely exclusive, then one has reinstated the sort of consistency conditions that lead to paradox in the first place, the sort of conditions that the dialetheic paraconsistent approach councils against, against a difference between 'not-(A is true)' and 'not-(A) is true'.

With a fully paraconsistent metatheory there is no in-principle problem with inconsistent set membership. Going non-classical at the meta-level, indeed, one could add to the above semantic conditions the stipulation that truth and falsity *contradict*: adopting both EXCLUSION and EXHAUSTION gives natural 'contradictory' interaction between t and f. On this track, 'excludes' itself can be inconsistent.

Similarly for $\mathscr{I}^+(F)$ and $\mathscr{I}^-(F)$, the intended picture of the standard presentation is



If these are classical, then while for some a it is allowed that

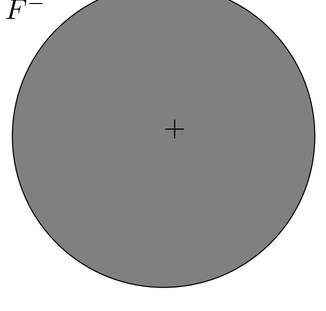
$$a \in \mathscr{I}^+(F) \cap \mathscr{I}^-(F)$$

one of the following would still have to go:
Exhaustion. If $a \notin \mathscr{I}^+(F)$ then $a \in \mathscr{I}^-(F)$.
Exclusion. If $a \in \mathscr{I}^-(F)$ then $a \notin \mathscr{I}^+(F)$.
Else there are cases where $a \in \mathscr{I}^+(F)$ and $a \notin \mathscr{I}^+(F)$.

Again, though, arguably this sort of external contradiction is just what a dialetheist means, to preserve the contradictory interactions,

$$\forall x(x \in \mathscr{I}^+(F) \text{ iff } x \notin \mathscr{I}^-(F))$$
$$\forall x(x \in \mathscr{I}^-(F) \text{ iff } x \notin \mathscr{I}^+(F))$$

The picture returns to



Allowing this back in (or never having pushed it out) means there will be significant 'noise' in the models; but part of the aim was to listen to the noise.[6]

2.3. Models Redux: Listen to the Noise

We now reformulate the notion of a model, using a background paraconsistent logic and set theory, as in Section 1.5. This is done, principally, by dropping the positive and negative valences on interpretations — which are not so covert ways of reintroducing consistency — and instead returning to the simpler idea that truth and falsity are dual.

DEFINITION 3 (Glutty Models). For $\mathcal{M} = \langle D, \mathscr{I} \rangle$, if F is an n-place predicate, then we say simply

$$\mathscr{I}(F) \subseteq \{\langle d_1, \ldots, d_n \rangle : d_1, \ldots, d_n \in D\}$$

For identity, everything is self-identical, $\mathscr{I}(=) = \{\langle x, x \rangle : x \in D\}$.

Since $A \vee \neg A$ and $\neg(A \mathbin{\&} \neg A)$ are theorems of our background logic, we have automatically that both exhaustion and exclusion — paraconsistently understood — hold:

$$\forall x \in D(x \in \mathscr{I}(F) \vee x \notin \mathscr{I}(F))$$
$$\neg\exists x \in D(x \in \mathscr{I}(F) \mathbin{\&} x \notin \mathscr{I}(F))$$

(These *are* formulated using restricted quantification, an issue we will note in Section 5.1 below.) Similarly with identity, nothing is not identical to itself, $\neg\exists x \in D^2(x \notin \mathscr{I}(=))$, even if also sometimes some things *are* non-self-identical. Rather than give independent truth and untruth clauses, by restoring the connection, duality means that one already determines the other (even if also they can diverge).

Then the base conditions on semantics look very familiar indeed:

[6] Of course, one may fairly ask (as a referee does) whether the picture is really so felicitous. In inconsistent instances, surely, the paraconsistent model theory clearly *differs* (in some sense) from the classical one. Perhaps we should add that there is still space in the picture for some inconsistency along the boundary — but in the same breath, we should reiterate that *there are no true contradictions*, since $\neg(A \wedge \neg A)$ is valid in our background logic. This is a question of the ORDINARINESS HYPOTHESIS in practice; we will indeed see some problems in the next section. See also (Weber et al, 2016) and (Weber, 2021, ch. 10).

DEFINITION 4. A *glutty pre-valuation* v is such that

$$\begin{array}{lll} \mathsf{t} \in v(F(t_0, \ldots, t_n) & \text{iff} & \langle \mathscr{I}(t_0), \ldots, \mathscr{I}(t_n)\rangle \in \mathscr{I}(F) \\ \mathsf{f} \in v(F(t_0, \ldots, t_n) & \text{iff} & \langle \mathscr{I}(t_0), \ldots, \mathscr{I}(t_n)\rangle \notin \mathscr{I}(F) \end{array}$$

The other semantic conditions are as in Def. 2 above, plus

$$\begin{array}{lll} \mathsf{t} \in v(A) & \text{iff} & \mathsf{f} \notin v(A) \\ \mathsf{f} \in v(A) & \text{iff} & \mathsf{t} \notin v(A) \end{array}$$

The semantics reflect both the law of excluded middle and the law of non-contradiction,

$$(\mathsf{t} \in v(A) \vee \mathsf{f} \in v(A)), \quad \neg(\mathsf{t} \in v(A) \;\&\; \mathsf{f} \in v(A))$$

both valid in the background and foreground logic. This, I would submit, multiplies contradictions as necessary.[7]

It is now important to see that, contra remark 2,

Remark 3. With a paraconsistent metatheory, two valued relational semantics are *not* equivalent to a three valued functional presentation (Priest, 2008, p. 151, footnote 5).

To see this, suppose we have a binary relation

$$R \subseteq \{a\} \times \{0, 1\}$$

There are three cases (omitting the empty case), which would *seem* to correspond precisely to three states:

$$\begin{array}{rll} R(a, 0) \text{ and } \neg R(a, 1) & \text{`iff'} & R(a) = \{0\} \\ R(a, 1) \text{ and } \neg R(a, 0) & \text{`iff'} & R(a) = \{1\} \\ R(a, 0) \text{ and } R(a, 1) & \text{`iff'} & R(a) = \{0, 1\} \end{array}$$

But the iffs are in scare quotes, because if $\neg$ is negation in a paraconsistent language, then $\neg R(a, 1)$ does not rule out $R(a, 1)$. Whereas $R = \{\langle a, 0\rangle\}$ *does* rule out $\langle a, 1\rangle \in R$, by the law of ordered pairs: if $\langle a, 1\rangle \in \{\langle a, 0\rangle\}$ then $\langle a, 1\rangle = \langle a, 0\rangle$ and then $0 = 1$. Assuming that $0 = 1$ is unacceptable (by the PARACONSISTENCY THESIS; cf. Section 3.3

[7] There may be a connection here with Priest's work on *hypercontradictions* (Priest, 1984), as a referee points out, though note that study is situated in classical metatheory and so is taking a different perspective.

below) then the two formalisms — three valued functions and two valued relations — are not equivalent.[8]

Making these arrangements on truth is not so difficult. Matters become more complicated when we apply this account to validity and proof.

3. Valid and Invalid

3.1. Truth in a model

DEFINITION 5 (Semantic Validity). A sentence is *true in a model* $\mathcal{M} \Vdash A$, when $\mathsf{t} \in v_{\mathcal{M}}(A)$. A sentence is a *theorem* when it is true in every model. An argument is *semantically valid* $A_0, \ldots, A_n \vDash B$ iff for every model, if $\mathcal{M} \Vdash A_0, \ldots, A_n$ then $\mathcal{M} \Vdash B$.

An argument is valid iff all assignments making premises true make the conclusion true. An argument is *invalid* iff there is an assignment making the premises true but the conclusion untrue.

As before, it is tempting to distinguish 'positive' and 'negative' clauses, e.g., with notation like $\mathcal{M} \Vdash^{+} A$ and $\mathcal{M} \Vdash^{-} A$. But instead the dialetheist can follow the simpler path already before us and just say

$$\mathcal{M} \Vdash A \text{ iff } \mathsf{t} \in v_{\mathcal{M}}(A) \qquad \mathcal{M} \nVdash A \text{ iff } \mathsf{t} \notin v_{\mathcal{M}}(A)$$
$$\mathcal{M} \Vdash \neg A \text{ iff } \mathsf{f} \in v_{\mathcal{M}}(A) \qquad \mathcal{M} \nVdash \neg A \text{ iff } f \notin v_{\mathcal{M}}(A)$$

So $\mathcal{M} \Vdash A$ or $\mathcal{M} \Vdash \neg A$, and not both and sometimes also both and that is that.

[8] To prove this (classically (which I've urged we should not feel obliged to do)), we could use a three-valued classical model. Let ν be a function from the language to $\{\mathsf{t}, b, \mathsf{f}\}$. Then

$$\begin{aligned} \nu(R(a,0) \vee R(a,1)) &= \mathsf{t} \\ \nu(\neg R(a,1)) &= b \\ \nu(R(a,0)) &= \mathsf{f} \end{aligned}$$

Thus, if we have soundness, the move from $1 \notin R(a)$ to $R(a) = \{0\}$ is not valid. See (Priest, 2006, p. 287).

3.2. The Invalidity Problem

Except that may not be that. In the language of LP, for any propositional atom p, consider the assignment v such that

$$\langle p, z\rangle \in v \Leftrightarrow (z = \mathsf{f} \vee z = \mathsf{t})$$

which makes every sentence both true and false: Then

$$\begin{array}{ll} \langle p, \mathsf{t}\rangle \in v & \langle \neg p, \mathsf{t}\rangle \in v \\ \langle p, \mathsf{f}\rangle \in v & \langle \neg p, \mathsf{f}\rangle \in v \end{array}$$

The existence of such a valuation is not news; its existence is an exercise in (Priest, 2008, 8.10, problem 5).

If we endorse EXCLUSION and EXHAUSTION, though, then this valuation does not merely make everything true and false; it makes everything true and untrue. We could then argue about whether or not to endorse EXCLUSION in particular — but it would be a distraction, since we can arrive at the same problem more indirectly. Let $\mathfrak{p}$ be some true contradiction. Consider the assignment $v^{\mathfrak{p}}$ such that

$$\langle p, z\rangle \in v^{\mathfrak{p}} \Leftrightarrow (z = \mathsf{f} \vee z = \mathsf{t}) \ \&\ \mathfrak{p}$$

Then since a conjunction with a false conjunct is false,

$$\begin{array}{ll} \langle p, \mathsf{t}\rangle \in v^{\mathfrak{p}} & \langle p, \mathsf{t}\rangle \notin v^{\mathfrak{p}} \\ \langle p, \mathsf{f}\rangle \in v^{\mathfrak{p}} & \langle p, \mathsf{f}\rangle \notin v^{\mathfrak{p}} \end{array}$$

Then some assignment makes *any* A true and untrue, and so $A \nvDash A$ because there is a way the premise can hold without the conclusion, for all A. The argument from A to itself is invalid (as well as valid) and indeed *every* argument is invalid; see (Batens, 2019; Young, 2019; Priest, 2020).

It remains the case that only some — not all — arguments are still valid, but this still seems bad. What is the point, in the end, of a dialetheist saying that *ex false quodlibet* is invalid if *everything* is invalid?

3.3. A methodological note

How bad is the invalidity problem for a dialetheist?

Dialetheism contradicts itself — dialetheism itself is both true and false, according to itself (Priest, 1979) — and rightly so. This is not however a ‘get out of jail free’ card that allows us, for example, to use some paraconsistently invalid notions in the background framework and then excuse doing so as just more acceptable inconsistency. That itself would be a bad methodological slide, of the form “since I allow and admit some contradictions somewhere, I can allow any contradiction anywhere.” That is self-defeating, or *incoherent*.

Paraconsistency is a tool that comes in after we’ve already made a commitment to not be self-defeating. You wouldn’t say “I am consistent because I use classical logic”; rather you say “I use classical logic because I am consistent.” Classical logic does not guarantee consistency. It expresses a commitment to consistency. So too for paraconsistency: we begin with a commitment to coherence, rather than expect our formalism alone to guarantee it. The PARACONSISTENCY THESIS from Section 1.4 is there to say that something is still missing from our account.

3.4. A solution

To prevent *all* arguments from being rendered invalid, we assume at least one proposition $\bot$, that is *absolutely false*, never in any way true. This is in our background logic, as the axiom $\bot \Rightarrow A$. As in Section 1.4 above, paraconsistency can be taken as the thesis that at least one proposition is utterly false, even if some contradictions are true — although now we are amping that up a little to say that some proposition is *necessarily* false (not at all true, in all interpretations). One such proposition is $\mathsf{t} = \mathsf{f}$ since if this were true then there is only one truth value and hence no point in ever asking what truth value any sentence has. Similarly for $0 = 1$. Another such proposition is $\forall x \forall y x \in y$ since then every object would have every property, which would mean $\mathsf{t} = \mathsf{f}$ again. We listen to the noise; we don’t blow out our eardrums.

The following definition makes this official.

DEFINITION 6. A *valuation* is a pre-valuation (def 4) such that $\mathsf{t} \in v(\bot)$ iff $\bot$.

The ‘$\bot$’ on the right is a condition in the meta-theory, i.e., a valuation is such that if it assigned truth to $\bot$ in a model then $\bot$ would be true *in real life*, which it is not and cannot be, on pain of absurdity. Under this definition there is no coherent trivial interpretation.[9]

Whether there are further problems to face here, we return to in (Weber, 2021, ch. 10) and in Section 5.1 below.

4. Provable and Unprovable

From semantics we turn to *derivations*. An argument is proof theoretically valid iff it is derivable in a proof system, e.g., if there is a sequence leading from the premises to the conclusion by means of valid rules. An argument is invalid iff there is no such sequence. One might think that the abstractions of model theoretic semantics make imaginings about dialetheias fairly easy; whereas when it comes to the concrete, down-to-earth business of constructing proofs out of fixed and finite ink marks, such story-telling becomes much harder. Let us see.

4.1. Soundness and Completeness

Soundness and completeness theorems are usually invoked so that derivation failures correspond to counterexamples. Here we can do as follows.

Here is a natural deduction system for LP (cf. Priest, 2002b, p. 309).

$$\frac{A \quad B}{A \wedge B} \qquad \frac{A \wedge B}{A} \qquad \frac{A \wedge B}{B} \qquad \frac{A}{A \vee B} \qquad \frac{B}{A \vee B}$$

$$\frac{A \vee B \quad \begin{matrix}\overline{A}\\ \vdots \\ C\end{matrix} \quad \begin{matrix}\overline{B}\\ \vdots \\ C\end{matrix}}{C} \qquad \frac{}{A \vee \neg A} \qquad \frac{\neg\neg A}{A} \qquad \frac{A}{\neg\neg A} \qquad \frac{\neg(A \wedge B)}{\neg A \vee \neg B}$$

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

[9] The same result might be achieved in a more fine grained way, by specifying for each interpretation at least one sentence that is not satisfied, but it could be a different sentence in different models, rather than a single sentence $\perp$ that cuts across all of them. Since $\perp$ is already available, though, the approach here is simpler. Again see (Badia, 2016) for the role of $\perp$ in controlling the structure of models.

DEFINITION 7 (Proof theoretic validity). An argument from $A_0, \ldots, A_n$ to B is *valid* $A_0, \ldots, A_n \vdash B$ if and only if there is a derivation of B from $\{A_0, \ldots, A_n\}$ using the above rules.

Using this we can establish soundness and (weak) completeness correspondence between the natural deduction system and the semantics of def. 2 with respect to LP, namely that definitions 5 and 7 are equivalent. This is a modification of (Weber et al, 2016; Badia et al., 2022; Omori and Weber, 2023).

THEOREM 4 (Soundness). *If* $A_0, \ldots, A_n \vdash A$ *then* $A_0, \ldots, A_n \vDash A$.

PROOF. Each of the rules of the natural deduction system are semantically valid. For example, $\neg(A \vee B) \vdash \neg A \vee \neg B$ is truth preserving. Proof: using the LP semantics, suppose $\mathsf{t} \in v(\neg(A \vee B))$. Then $\mathsf{f} \in v((A \vee B))$ and then $\mathsf{f} \in v(A)$ and $\mathsf{f} \in v(B)$. Then $\mathsf{t} \in (\neg A)$ and $\mathsf{t} \in v(\neg B)$, and so $\mathsf{t} \in v(\neg A \wedge \neg B)$ as required. The other cases are similar. ⊣

Proving (weak) completeness is by a constructive method due to Kalmár in 1935. Notation: for an LP-interpretation v, let

$$\mathcal{A} = \{p_i : \mathsf{t} \in v(p_i)\} \cup \{\neg p_i : \mathsf{f} \in v(p_i)\}$$

where $\{p_0, \ldots, p_n\}$ is the set of all the propositional atoms occuring in A.

THEOREM 5 (Completeness). *If* $\vDash A$ *then* $\vdash A$.

PROOF. First we show that, for any LP-interpretation v,

1. if $\mathsf{t} \in v(A)$ then $\mathcal{A} \vdash A$
2. if $\mathsf{f} \in v(A)$ then $\mathcal{A} \vdash \neg A$

The proof is by induction.[10] If A is a propositional atom then the result is immediate. For complex formulas, to illustrate, consider conjunction. For (1), if $\mathsf{t} \in v(A \wedge B)$ then $\mathsf{t} \in v(A)$ and $\mathsf{t} \in v(B)$. By induction hypothesis, $\mathcal{A} \vdash A$ and $\mathcal{B} \vdash B$. So then $\mathcal{A}, \mathcal{B} \vdash A \wedge B$. For (2), if $\mathsf{f} \in v(A \wedge B)$ then $\mathsf{f} \in v(A)$ or $\mathsf{f} \in v(B)$. Then $\mathcal{A} \vdash \neg A$ or $\mathcal{B} \vdash \neg B$. Then $\mathcal{A}, \mathcal{B} \vdash \neg A \vee \neg B$, and then $\mathcal{A}, \mathcal{B} \vdash \neg(A \wedge B)$. Other cases are similar.

Now suppose $\vDash A$, that is, $\mathsf{t} \in v(A)$ for every v. There are finitely many assignments $v_0, \ldots, v_m$ of different values — either f or t — to the

[10] Recall in Section 1.5 above that we assume some limited form of mathematical induction, a practice which seems essential and is continued in (Badia et al., 2022) — but which is not completely unproblematic. See (Weber, 2021, ch. 6).

propositional atoms in A. For each such v_i build the finite set $\mathcal{A}_{v_i}$. Then as we just saw, $\mathcal{A}_{v_i} \vdash A$, for each $i \leqslant m$. Then these $\mathcal{A}_{v_i}$ can be combined into a provable theorem by excluded middle; this provable theorem proves A; and the result follows by transitivity of consequence. $\dashv$

Note that the 'if / then' in the statements of both soundness and completeness is the $\Rightarrow$ of our metatheory — which does not contrapose. That is why we gave direct arguments for each, rather than the nowadays more standard proofs showing that *counterexamples* correspond to an *absence* of a derivation. In this context, the existence of a counterexample may *not* quite always amount to the absence of a derivation; it may amount to a derivation that is not a derivation (and so in *that* sense an 'absence'), in a sense we now explore.

4.2. Inconsistent Proofs

If a proof is a sequence of propositions arrived at by step-by-step application of valid rules, it is perhaps harder than truth or validity to see how derivations might be inconsistent (Shairo, 2002). Nevertheless, there is reason to think the proof relation is inconsistent (Priest, 1979; Routley, 1979; Berto, 2007; Weber, 2022). If

$G =$ 'this sentence is unprovable'

is both true and false there is a sentence which is both provable and unprovable, which means that there is both a proof of it and there is no proof of it. What does that *look* like? What is an inconsistent proof?

It is tempting to begin to imagine some sort of strange or paradoxical objects, like the tape of a Turing Machine wrapped into a Möbius strip, or a device built from some exotic quantum materials. A guiding methodological idea — the ORDINARINESS HYPOTHESIS — is to demystify the problem, and adopt the attitude that what we are seeking is new *not* because of some surprising technology, but because of new ways of *thinking* about regular technology.

A proof is a sequence. If there is a proof of G then we know what that looks like: $\langle A_0, \ldots, A_n, G\rangle$. If there is no proof of G then there is no sequence ending with G with certain properties. How can that be, especially if we already accept that there is at least one such sequence?

Let $\langle a_0, a_1\rangle$ be a (short) sequence. Set theoretically this is broken down to be

$$\begin{aligned}\{\{a_0\},\{a_0,a_1\}\} &= \{\{x : x = a_0\},\{x : x = a_0 \vee x = a_1\}\}\\ &= \{y : y = \{x : x = a_0\} \vee y = \{x : x = a_0 \vee x = a_1\}\}\end{aligned}$$

using the standard definitions of singletons and pairs, underwriting the law of ordered pairs, as in Section 1.5 above.

Now suppose $a_0 \neq a_0$. (Don't panic; just suppose. We still know that $a_0 = a_0$, because it is an axiom that everything is self-identical, and we could even prove it to be so if we defined identity in terms of a biconditional so that $x = y \Leftrightarrow \forall z(x \in z \Leftrightarrow y \in z)$. Still, an object like the Russell set will, additionally, be non-self-identical, as in remark 1 above.) Then $a_0 \in \{a_0\}$, and $a_0 \notin \{a_0\}$, so by Axiom 2, $\{a_0\} \neq \{a_0\}$. Similarly, $\langle a_0, a_1\rangle = \langle a_0, a_1\rangle \neq \langle a_0, a_1\rangle$.

So too for any tuple $\langle \ldots, a_0, \ldots\rangle$ containing a_0. This more general fact is proved with the help of the following lemma.

LEMMA 6 (Non-self-identity). *If $a \neq a$ then $\forall y(a \neq y)$.*

PROOF. Suppose $y = a$. If a has property $a \neq a$ then $y \neq y$ too, by substitution. ⊣

PROPOSITION 7. *If $Y = \{a_0, \ldots, a_n\}$ and $a_i \neq a_i$ for some $i \in \{0, \ldots, n\}$ then $Y \neq Y$.*

PROOF. For

$$a_i \notin Y \Leftrightarrow a_i \neq a_0 \;\&\cdots\&\; a_i \neq a_n$$

by axiom 2. So $a_i \notin Y$, by the previous lemma. But also $a_i \in Y$. So $Y \neq Y$. ⊣

Note that we are *not* proving that in general $x \neq x \;\&\; x \in y \Rightarrow y \neq y$. This only holds when y is extensionally listed out using identity. Some sets can't be so listed, like $\mathcal{V} = \{x : \exists y(x \in y)\}$. Nevertheless we can lift the result to tuples.

THEOREM 8. *If $a_i \neq a_i$ for some $0 \leqslant i \leqslant n$, then $\neg\exists Y(Y = \langle a_0, \ldots, a_n\rangle)$.*

PROOF. Let $a \neq a$, and $Y = \langle a_0, \ldots, a_n\rangle$, and $a = a_0 \vee \ldots \vee a = a_n$. As this is for any n, we show $a \notin Y$ by induction. For the base case: $a \notin \langle a_0\rangle = \{a_0\}$. For the induction step: suppose $a \notin \langle a_0, \ldots, a_k\rangle$ for $0 \leqslant k < n$. Then since $\langle a_0, \ldots, a_{k+1}\rangle = \langle\langle a_0, \ldots, a_k\rangle, a_{k+1}\rangle$, and we know $a \neq$

a_{k+1}, we have $a \notin \langle a_0, \ldots, a_{k+1} \rangle$. Thus $\langle a_0, \ldots, a_n \rangle \neq \langle a_0, \ldots, a_n \rangle$, and so ex hypothesi $\forall Y(Y \neq \langle a_0, \ldots, a_n \rangle)$, and the conclusion follows. ⊣

The intended application here is that if a sequence of formulas $\langle A_0, \ldots, A_n, G \rangle$ includes an inconsistent element, say $G \neq G$, then the sequence is inconsistent $\langle A_0, \ldots, A_n, G \rangle \neq \langle A_0, \ldots, A_n, G \rangle$ and so it is *not* a sequence: $\neg \exists Y(Y$ is a sequence ending in $G)$.

This begins to answer the question of what it could mean for something to be both provable and not: there is a derivation, and also there is no derivation, because the derivation contains a contradictory element. An inconsistent proof — a proof that is a non-proof — looks like a proof, ending with an inconsistent object.[11]

Priest (Priest, 2006, p. 242) outlines how the Gödel code of the Gödel sentence would correspond to an inconsistent *number*, in a model of inconsistent arithmetic. Here we are taking a more direct approach by describing the object itself as inconsistent, not just its coding.

What, then, does an inconsistent object look like? ORDINARINESS already tells us the basics of the answer — it looks like an *object* — but going into that discussion will have to wait for another day.

5. Conclusion

5.1. Assessment

The aim of this paper has been to present a fairly extreme approach to paraconsistent (meta)theory, in order to evaluate its philosophical and technical tenability.

5.1.1. What has been done

There is good philosophical support for such an approach, and hence for facing the challenges it may throw up. Guided by the ORDINARINESS

[11] A referee notes that, given some proof $\langle A_0, \ldots, A_m \rangle$ we could add an irrelevant 'detour' along the way — informally, a few extra sentences does not stop a proof from being a proof; but such a detour could include an inconsistent object, so there is prima facie always a way to make a proof inconsistent. This is not too different from the general situation of being able to 'inconsistentize' most anything, as we saw with validity and the counterexample problem. Turning a proof into an inconsistent proof, note, does not ipso facto make the original proof inconsistent. But a fuller account of inconsistent proof theory is called for.

and non-classicality hypotheses, we get a picture that on the surface level is nearly indistinguishable from familiar presentations of *classical* logic: with truth and falsity contradicting, conditions on models can be stated easily and naturally. By respecting the Paraconsistency thesis we can see why the 'invalidity problem' can be avoided and can go on to establish soundness and completeness results for the extensional fragment of the language, as well as outline the start of a paraconsistent computability picture. What I hope this picture presents is, to paraphrase Schopenhauer, that we are not trying to see new exotic objects that no one has yet seen, but rather to think in new ways about the old objects that everybody sees.

5.1.2. What is still to do

There are, though, some very significant technical challenges to face on this sort of radical approach. Let me name a few here, as open problems.

We have avoided saying anything about the conditional operator $\Rightarrow$, which is not truth functional and not particularly amenable to easy treatment. In general $A \Rightarrow B$ is not equivalent to $\neg(A \mathbin{\&} \neg B)$, which overall is good, but does result in some awkwardness, e.g., it feeds into problems with duality and restricted quantifiers. In Section 2.3 above we stated as duals expressions of the form

$$\forall x \in D(Ax \vee \neg Ax), \quad \neg\exists x \in D(Ax \mathbin{\&} \neg Ax)$$

These are not dual though if $\forall x \in D(Ax), \exists x \in D(Ax)$ are defined as $\forall x(x \in D \Rightarrow Ax), \exists x(x \in D \mathbin{\&} Ax)$, respectively. This becomes particularly plangent when we come to the notion of validity itself—usually stated in conditional terms: an argument is valid iff if every model of the premises is a model of the conclusion, and invalid when some model of the premises is not a model of the conclusion. But if these clauses are $\forall \Rightarrow$, $\exists\&$ sentences respectively, then there is a gap between them. Perhaps such core notions, in the end, have faced such radical testing that they need to be revisited in paraconsistent light.

A recurring theme of the technical challenge is that paraconsistency can be very good at global 'absolute generality', so the existence of a set of all sets or unrestricted universal quantification are unproblematic, but then struggles with local restrictions, so the existence of successor ordinals (a still unsolved problem from (Weber, 2010), cf. (Weber, 2021, p. 190)) or restricted quantification becomes more difficult. Indeed, we have avoided details about very basic things, like numbers in finite arithmetic, that are used to track step-by-step processes. There are paraconsistent theories of arithmetic that would seem to do most of what one wants here (see. e.g., Weber, 2021, ch.6), but there are also many other better studied paraconsistent arithmetics with inconsistent models that leave the required properties underdetermined (Ferguson, 2019; Tedder, 2021), cf. (Weber, 2022, §3). Most of the hard technical work remains to be done.

5.1.3. A further problem

Let me admit here a problem that has a technical solution but remains, I think, philosophically vexing. As flagged in Section 1.5, once we admit the $\bot$ particle as absolutely unacceptable, we can define a kind of 'or-else negation' $A \Rightarrow \bot$ that allows one to express 'this sentence is false, on pain of absurdity'. Formally, this amounts to the Curry sentence, $C \Leftrightarrow (C \Rightarrow \bot)$. The logical system used in our metatheory is designed to handle Curry sentences, by disavowing contraction on all levels. We have mathematical reassurance that the resulting system is not trivial. But there is something philosophically discordant about handling the Curry problem this way. The whole dialetheic paraconsistent approach as I've pitched it is about accepting the impossible fixed point knot in the foundations of logic as a true contradiction. We cannot, it seems, uncritically apply the same methodology to Curry.

Criticism of dialetheic paraconsistency along these lines — about Curry's paradox, revenge, and how it relates to Priest's *principle of uniform solution* (that the same type of problem should have the same type of solution) — have been aired repeatedly, including in (Beall, 2014; Burgis and Bueno, 2019); cf. (Murzi and Rossi, 2020). I think there is something to say in reply here — that dropping contraction makes dialetheic sense, if a true sentence A is different from the (very same) false sentence A, so $A \,\&\, A$ is as far from A as $A \,\&\, B$. And then the revealed answer to Curry is that

> if this sentence is true *and* this sentence is true, then $\bot$

is deeply different from

> if this sentence is true, then $\bot$

and each would require an additional iteration of 'this sentence is true' to be proven so. One might worry that, if this can be made sense of, it is still a hierarchical solution of the sort dialetheism is at pains to avoid. One might reply that just because some hierarchies are untenable does not make all iterative structures bad. In any case, there seems to be both philosophical and technical work here to do. We've dealt with truth and falsity, but the line between coherence and absurdity remains unresolved.

5.1.4. A way forward?

Finally, a fair question is whether much of what I've been suggesting *can* be done.[12] Of course, that may seem like an odd question, since a great deal of ingenious technical work already *has* been done, by a great many dedicated researchers going back at least to the 1950s (surveyed in (Weber, 2022)). Most of that work, though, is based eventually on classicality or some other form of conventional reassurance, and not in the fully radical paradigm I've been outlining here and elsewhere. The radical paraconsistent project — can *that* be done? If it can, it will take not only some bravery and perseverance, but also a willingness to think creatively, to reformulate not only the answers, but the questions themselves.

5.2. Ordinary Non-classicality

Let me end a rather programmatic paper on a programmatic note.

If the world as we know it is consistent and behaves according to classical logic, there will never be a satisfactory answer to what the 'paraconsistent' bits of the world are like. There are no such bits. If the world as we know it behaves according to paraconsistent logic then there is already a simple answer to what the paraconsistent bit of the world is like. It is the world. If there is a project to do here, the challenge is not to add new truth values, or new modalities, or more valences, or the like, to classical logic; it is to start to look at the non-classical as ordinary, and the ordinary as non-classical. And then the real, hard work can begin.

Acknowledgements. Thanks to Marek Nasieniewski and the organizers of the Sixth World Congress of Paraconsistency in Toruń, Poland, and to the audience there, as well as audiences at Utrecht, the ILLC Amsterdam, Ostrava, Louvain-la-Neuve, Kobe, and Ohio State. Thanks to

[12] Already in 1982, Arruda and Batens say: "There can be no doubt that the expectations of paraconsistent logicians with respect to set theory did not come true" (Arruda and Batens, 1982, p. 132).

the editors of this special issue. Further thanks alphabetically to Sara Ayhan, Guillermo Badia, Jc Beall, Libor Behounek, Fernando Cano-Jorge, Caitlin Canonica, Colin Caret, Walter Carnielli, Marcelo Coniglio, Michael De, Sergey Drobyshevich, Thomas Ferguson, Dustin Gooßens, Wessel Kroon, Nils Kürbis, Franci Mangraviti, Satoru Niki, Hrafn Oddsson, Grigory Olkhovikov, Hitoshi Omori, Graham Priest, Sam Sanders, Fatima Scha, Daniel Skurt, Andrew Tedder, Ebru Tekin, Peter Verdée, Albert Visser, and Heinrich Wansing. And probably others, too. For support, thanks to the Alexander von Humboldt Foundation, Ruhr-Universität Bochum, and the Royal Society Te Apārangi of New Zealand.

References

Arruda, A. I., and D. Batens, 1982, "Russell's set versus the universal set in paraconsistent set theory", *Logique et Analyse*, 25 (8): 121–133.

Badia, G., 2016, "The relevant fragment of first order logic", *Review of Symbolic Logic*, 9 (1): 143–166. DOI: 10.1017/S1755020315000313

Badia, G., and Z. Weber, 2019, "A substructural logic for inconsistent mathematics", pages 155–176 in A. Reiger and G. Young (eds.), *Dialetheism and its Applications*, Springer. DOI: 10.1007/978-3-030-30221-4_9

Badia, G., Z. Weber, and P. Girard, 2022, "Paraconsistent metatheory: new proofs with old tools", *Journal of Philosophical Logic*, 51 (4): 825–856. DOI: 10.1007/s10992-022-09651-x

Batens, D., 2019, "Looting liars masking models", pages 139–164 in C. Baskent and T. Ferguson (eds.), *Graham Priest on Dialetheism and Paraconsistency*, Springer. DOI: 10.1007/978-3-030-25365-3_8

Batens, B., 2020, "Adaptive Fregean set theory", *Studia Logica*, 108 (5): 903–939. DOI: 10.1007/s11225-019-09882-1

Beall, Jc, 2009, *Spandrels of Truth*, Oxford University Press. DOI: 10.1093/acprof:oso/9780199268733.001.0001

Beall, Jc, 2014, "End of inclosure", *Mind*, 123 (491): 829–849. DOI: 10.1093/mind/fzu075

Beall, JC, R. T. Brady, A. P. Hazen, G. Priest, and G. Restall, 2006, "Relevant restricted quantification", *Journal of Philosophical Logic*, 35 (6): 587–598. DOI: 10.1007/s10992-005-9008-5

Beall, JC, and M. Colyvan, 2001, "Looking for contradictions", *Australasian Journal of Philosophy*, 79: 564–569.

Beall, Jc, and S. A. Logan, 2017, *Logic: The Basics*, Routledge. DOI: 10.4324/9781315723655

Beall, JC, and G. Restall, 2005, *Logical Pluralism*, Oxford University Press. DOI: 10.1093/acprof:oso/9780199288403.001.0001

Brady, R., 2006, *Universal Logic*, CSLI, Stanford.

Berto, F., 2007, *How to Sell a Contradiction*, vol. 6 of Studies in Logic, College Publications.

Burgis, B., and O. Bueno, 2019, "Liars with curry: Dialetheism and the prospects for a uniform solution", in A. Reiger and G. Young (eds.), *Dialetheism and its Applications*, Springer.

Cantini, A., 2003, "The undecidability of Grĭsin's set theory", *Studia Logica*, 74 (3): 345–368.

Carnielli, W., and M. E. Coniglio, 2016, *Paraconsistent Logic: Consistency, Contradiction and Negation*, Springer. DOI: 10.1007/978-3-319-33205-5

Ferguson, T. M., 2019, "Inconsistent models (and infinite models) for arithmetics with constructible falsity", *Logic and Logical Philosophy*, 28 (3): 389–407. DOI: 10.12775/LLP.2018.011

Ficara, E., 2021, "The birth of dialetheism", *History and Philosophy of Logic*, 42 (3): 281–296.

Field, H., 2020, "Properties, propositions, and conditionals", *Australasian Philosophical Review*, 4 (2): 112–146. DOI: 10.1080/24740500.2021.1886687

Istre, E., 2017, "Normalized naive set theory", PhD Thesis, University of Canterbury.

Libert, T., 2003, "ZF and the axiom of choice in some paraconsistent set theories", *Logic and Logical Philosophy*, 11: 91–114. DOI: 10.12775/LLP.2003.005

Martinez, S. J., 2021, "Algebra-valued models for lp-set theory", *Australasian Journal of Logic*, 18 (7).

Meadows, T., and Z. Weber, 2016, "Computation in non-classical foundations?", *Philosophers' Imprint*, 16: 1–17.

Murzi, J., and L. Rossi, 2020, "Generalized revenge", *Australasian Journal of Philosophy*, 98: 153–177.

Omori, H., and Z. Weber, 2019, "Just true? On the metatheory for paraconsistent truth", *Logique et Analyse*, 248: 415–433.

Omori, H., and Z. Weber, 2023, "A note on the expressive completeness of LP in a metatheory without negation", manuscript.

Ono, H., and Y. Komori, 1985, "Logic without the contraction rule", *Journal of Symbolic Logic*, 50 (1): 169–201. DOI: 10.2307/2273798

Petersen, U, 2000, "Logic without contraction as based on inclusion and unrestriced abstraction", *Studia Logica*, 64: 365–403.

Priest, G., 1979, "The logic of paradox", *Journal of Philosophical Logic*, 8 (1): 219–241. DOI: 10.1007/BF00258428

Priest, G., 1984, "Hypercontradictions", *Logique et Analyse*, 107: 237–243.

Priest, G., 2002b, *Beyond the Limits of Thought*, 2nd edition, Oxford University Press (first edition 1995, Cambridge University Press). DOI: 10.1093/acprof:oso/9780199254057.001.0001

Priest, G., 2002b, "Paraconsistent logic", pages 287–394 in D. M. Gabbay and F. Günthner (eds.), *Handbook of Philosophical Logic*, 2nd edition, Vol. 6, Kluwer. DOI: 10.1007/978-94-017-0460-1_4

Priest, G., 2006, *In Contradiction: A Study of the Transconsistent*, 2nd edition, Oxford University Press, Oxford. DOI: 10.1017/CBO9780511801174

Priest, G., 2008, *An Introduction to Non-Classical Logic*, 2nd edition, Cambridge.

Priest, G., 2020, "Metatheory and dialetheism", *Logical Investigations*, 26 (1): 48–59. DOI: 10.21146/2074-1472-2020-26-1-48-59

Priest, G., and R. Routley, 1983, *On Paraconsistency*, Research School of Social Sciences, Australian National University. Reprinted as introductory chapters in (Priest et al., 1989).

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

Restall, G., 1992, "A note on naïve set theory in *LP*", *Notre Dame Journal of Formal Logic*, 33 (3): 422–432. DOI: 10.1305/ndjfl/1093634406

Ripley, D., 2015, "Naive set theory and nontransitive logic", *Review of Symbolic Logic*, 8 (3): 553–571. DOI: 10.1017/S1755020314000501

Routley, R., 1977, *Ultralogic as Universal?*. First appeared in two parts in *The Relevance Logic Newsletter* 2(1): 51-90, January 1977 and 2(2):138-175, May 1977. Reprinted as appendix to *Exploring Meinong's Jungle and Beyond* 1980, pp. 892–962. New edition as *The Sylvan Jungle*, Vol. 4, edited by Z. Weber, Synthese Library, 2019.

Routley, R., 1979, “Dialectical logic, semantics and metamathematics”, *Erkenntnis*, 14 (3): 301–331. DOI: 10.1007/BF00174897

Routley, R., and R. K. Meyer, 1976, “Dialectical logic, classical logic and the consistency of the world”, *Studies in Soviet Thought*, 16: 1–25.

Shapiro, S., 2002, “Incompleteness and inconsistency”, *Mind*, 111 (444): 817–832. DOI: 10.1093/mind/111.444.817

Sylvan, R., and J. Copeland, 2000, “Computability is logic relative”, pages 189–199 in D. Hyde and G. Priest (eds.), *Sociative Logics and their Applications*, Ashgate.

Tedder, A., 2021, “On consistency and decidability in some paraconsistent arithmetics”, *Australasian Journal of Logic*, 18 (5): 473–502. DOI: 10.26686/ajl.v18i5.6921

Terui, K., 2004, “Light affine set theory: A naive set theory of polynomial time”, *Studia Logica*, 77: 9–40.

Weber, Z., 2010, “Transfinite numbers in paraconsistent set theory”, *Review of Symbolic Logic*, 3 (1): 71–92. DOI: 10.1017/S1755020309990281

Weber, Z., 2020, “Property identity and relevant conditionals”, *Australasian Philosophical Review*, 4 (2): 147–155. DOI: 10.1080/24740500.2021.1886688

Weber, Z., 2021, *Paradoxes and Inconsistent Mathematics*, Cambridge University Press. DOI: 10.1017/9781108993135

Weber, Z., 2022, *Paraconsistency in Mathematics*, Cambridge Elements, Cambridge University Press. DOI: 10.1017/9781108993968

Weber, Z., G. Badia, and P. Girard, 2016, “What is an inconsistent truth table?”, *Australasian Journal of Philosophy*, 94 (3): 533–548. DOI: 10.1080/00048402.2015.1093010

Young, G., 2019, “A revenge problem for dialetheism”, pages 21–46 in A. Reiger and G. Young (eds.), *Dialetheism and its Applications*, Springer. DOI: 10.1007/978-3-030-30221-4_2

ZACH WEBER
Department of Philosophy
University of Otago
Otago, New Zealand
zach.weber@otago.ac.nz