TRUTH AND PARADOX: A PHILOSOPHICAL SKETCH

JC Beall

1 Aims and structure of this essay

Because of the many surveys or otherwise readable discussions of truth and paradox, I have tried to take a slightly different (philosophical) perspective on the basic projects and issues. Moreover, I have tried to avoid focusing on the mathematics, leaving citations and 'further reading' to fill out the details; indeed, proofs are entirely left to the cited sources, almost all of which are readily available. The risk in pursuing these aims is that I will inevitably mischaracterize certain approaches, or at least mislead in various ways. That said, the riskiness is probably worth it. And when the risk of distortion—or being overly sketchy—is high, I try to flag alternative or primary sources to set things straight. (Unfortunately, this has resulted in an enormous amount of footnotes.) This essay is not intended as a representative survey of the many approaches or, per impossibility, issues that are significant in the literature on semantic (or logical) paradox; it aims merely to be a discussion of some of the issues and some of the approaches.

This essay is also abnormal in its chief focus, which is on *dtruth* (see §2), where dtruth is essentially intersubstitutable: $\mathsf{dT}\langle A\rangle$ and A are intersubstitutible in all (non-opaque) contexts.[1] Since the logic of dtruth cannot be classical (assuming the language to be suitably resourceful), theories of dtruth are accordingly non-classical. I focus largely on a few broad approaches to dtruth, namely, 'paracomplete' and 'paraconsistent'. On the other hand, many of the most popular approaches to truth-theoretic paradox are not theories of dtruth, but rather theories of some less transparent notion. I sketch—but, for space reasons, only very briefly sketch—a few such approaches in the last few sections. Further reading is given throughout, and a few remarks for further reading are given in §11. I should also mention that this essay focuses entirely on the semantic side of things. Some sources for axiomatic approaches are indicated in §11.

I have tried to avoid arguing or even commenting much on the adequacy (philosophical or otherwise) of the canvassed proposals. Instead, I try only to mention

[1] Note that, for the most part, I trust context to sort out use-mention. Except when it would impede readability, I write (e.g.) $\mathsf{T}\langle A\rangle$ instead of $\mathsf{T}(\langle A\rangle)$, despite T's being a unary predicate and, so, standardly written as $\mathsf{T}(x)$ or the like. I also leave out discussion of Gödel codes, etc., although this is certainly in the (mathematical) background.

Handbook of the Philosophy of Science. Volume 5: Philosophy of Logic
Volume editor: Dale Jacquette. Handbook editors: Dov M. Gabbay, Paul Thagard and John Woods.

one or two issues that are relevant to the adequacy of the proposals. I leave debate wide open, and otherwise to the cited material.

The basic structure of the essay is as follows. §2–§3 set up the main focus, namely, dtruth and dtruth-theoretic paradox, in particular, Liars. §4 covers a few different 'paracomplete' approaches to dtruth, beginning with a very informal, leisurely discussion of Kripke's non-classical (least fixed point, empty ground model) proposal. §5, in turn, focuses on 'paraconsistent' (and, in particular, 'dialetheic') approaches. §6 closes the discussion of dtruth with a brief discussion of validity and 'dtruth-preservation'. Turning away from dtruth, §7 briefly sketches a few standard parametric and contextualist approaches, including the 'indexicalist' idea, the 'quantifier-variability' proposal, and 'situational truth'. §8, in turn, briefly sketches the basic 'revision-theoretic' idea. The final two sections, §9 and §10, provide a few comments on related matters, with §9 mentioning 'set'-theoretic paradox and its relation to semantics, and §10 (very briefly) discussing 'revenge'.

Before turning to the discussion, I should mention the importance of truth-theoretic paradox, something often ignored or, perhaps, unappreciated by many philosophers working on the 'nature' of truth. (It goes without saying that the paradoxes are of fundamental importance in philosophical logic, but they are surprisingly not treated as pressing among philosophers, more generally, even among those who work on 'nature' questions.) By my lights, it is hard to be too invested in theories of the 'nature' of truth without having some grip on, some resolution of, the paradoxes. After all, suppose that one's 'nature' story turns on various key principles. It may well be that such principles cannot consistently—or, more broadly, even non-trivially—or even plausibly be maintained in a logic that accommodates the paradoxes. (This issue is particularly pressing when it comes to a suitable conditional, a topic discussed below.) Admittedly, the issues of 'nature' and logic are related, since one's 'nature' story will often motivate—if not dictate—a particular approach to the logic of truth. Still, it strikes me as methodologically misguided to plow forward on 'nature' questions without at least a steady eye on the truth-theoretic paradoxes. But, for space reasons, I will leave this matter—like many in this essay—open.

2 Dtruth and truth

There is an established usage of 'is true' according to which

$$\langle A \rangle \text{ is true}$$

and

$$A$$

are intersubstitutable in all transparent (non-opaque) contexts. Let us use 'dtrue' for the given usage, a (unary) predicate dT defined simply in terms of such intersubstitutivity.[2] If the language also has a conditional $\rightarrow$ such that every instance of $A \rightarrow A$ is dtrue, then the well-known dT-biconditionals are also, one and all, dtrue: $\mathsf{dT}\langle A\rangle \leftrightarrow A$.

We can think of 'dtrue' (or, as I'll say, dtruth) as a mere expressive device, one introduced only to circumvent practical, expressive problems. According to a common metaphor, we once spoke only the 'dtrue'-free fragment of our language. For the most part, the given fragment served our purposes well. We could say that Max is a cat, that Gödel and Tarski were independently ingenious, that there will never be cloned animals, and so on. Daily discourse, so long as it didn't generalize too much, worked well. But generalization is inevitable. Even in daily discourse, let alone science or logic, in general, the need arises to say (what, using dtruth, we say when we say) that all of So-and-so's assertions are dtrue (or that they're false, that is, that the negation of each of So-and-so's assertions is dtrue). Were we God (or even just beings with infinite time or capacities), we wouldn't need to use 'dtrue' in such generalizing contexts; we could simply assert each of So-and-so's assertions (or the negations thereof). But we're not, and so we introduced 'dtrue' to achieve the given sorts of expression. And *that*, and only that, is the job of 'dtrue' in our language.

The foregoing, of course, is but a metaphor, one that is common among so-called deflationists (or, in particular, disquotationalists). Debates rage over whether there is more to *truth* than dtruth, whether there's more to say about the 'nature' of truth than is given in the story about the *device* 'dtrue'. Deflationists, crudely put, maintain that we need not (indeed, ought not) pursue the 'nature' of truth: at bottom, the expressive device *dtrue* is all that we need acknowledge.[3] Non-deflationists maintain that there's more to truth than what is captured by our expressive device *dtrue*.

For present purposes, we can set aside the debate between deflationists and non-deflationists. Both sides can at least acknowledge the simple, 'merely disquotational' usage—what we're calling *dtruth*. The issue at hand concerns truth-theoretic paradox, which is accentuated in the case of dtruth. Given the inter-substitutivity of dtruth, any distinction between Bivalence (BIV) and the Law of Excluded Middle (LEM) collapses—at least with respect to what can be said *in* the given (object-) language. If, as is standardly assumed, falsity amounts to truth of negation (that A is false exactly if $\neg A$ is true), then 'dfalse' and 'not dtrue' are equivalent, that is, $\mathsf{dT}\langle\neg A\rangle$ and $\neg\mathsf{dT}\langle A\rangle$ are equivalent. In turn, $\mathsf{dT}\langle A\rangle \vee \neg\mathsf{dT}\langle A\rangle$ and $\mathsf{dT}\langle A\rangle \vee \mathsf{dT}\langle\neg A\rangle$ and $A \vee \neg A$ are equivalent, and so any 'distinction' between LEM and BIV collapses—at least *in* the language. Dtruth is entirely transparent.

[2] There is no aim to coin a neologism here, but rather only an attempt to clarify that we're talking about the 'merely expressive' usage of 'true'. I should note that, for purposes of pronunciation, 'dtrue' follows the conventions of Kaplan's 'dthat' [1989, p. 521, fn. 45].

[3] This does not preclude disquotationalists (or deflationists, generally) from acknowledging other 'truth-like' predicates that might do significant work (e.g., in semantics, science). The requirement on disquotationalists is that any other such predicates be ultimately definable in terms of dtruth and other logical devices.

3 Liars: broad picture and projects

A typical Liar is a sentence that says—or may be used to say—of itself only that it is false (or, equivalently in the case of dtruth, not true). Example:

$\surd$ The ticked sentence in §3 is false.

If the ticked sentence is true, it is false, and if false, true. Accordingly, if the ticked sentence is either true or false, it is apparently both true and false. That, in short, is the basic Liar paradox.[4]

Two related but distinct projects dominate the Liar-literature and work on semantic paradoxes, in general, at least among philosophical logicians concerned with modeling truth itself.

NTP. *Non-triviality Project*: Explain how, despite having a truth predicate (in our language, for our language) and Liar-sentences, our language is non-trivial.[5]

ECP. *Exhaustive Characterization Project*:[6] Explain how, if at all, we can truly characterize—specify the 'semantic status' of—all sentences of our language (in our language).

These projects reflect the core appearances that give rise to the Liar paradox (and its ilk). Semantic paradoxes arise, at least in part, from the appearance that we can 'exhaustively characterize' all sentences of our language in terms of 'semantically significant' predicates, and truly do as much in our language.[7]

Consider the classical picture according to which our semantically significant predicates are 'true' and 'false'. Our exhaustive characterization takes the form of bivalence.

CEC. *Classical*: Every sentence is either true or false.

[4]All of this can be formalized precisely. The achievements of Tarski [1936; 1944] (and, relatedly, Gödel) made plain that astonishingly few resources are needed to generate the Liar paradox: a sufficiently 'syntactically rich' language in which the T-biconditionals hold and for which classical logic holds.

[5]A *trivial language* (or theory) is one according to which everything is true. A *non-trivial language* is one that isn't trivial. [For convenience, I will often speak of the (non-) triviality of *truth* or 'true' (relative to some language) to mean the same as 'trivial language'. To say that truth—or 'true'—is trivial is to say that A is true (or A is in the extension of 'true') for *all* A in the given language.] Of course, most theorists are concerned with *consistency* (and, hence, non-triviality); however, 'glut-theorists' and paraconsistentists, in general, are concerned with (reasonable) non-trivial but (negation-) inconsistent languages/theories, and so 'non-triviality' is a more general term. (See §5 for paraconsistent approaches.)

[6]In his [1991], Patrick Grim uses 'complete' in the target sense of 'exhaustive'. Were it not for the already too many notions of 'complete' in logical literature, Grim's terminology would be quite appropriate here. [In particular, one can view the two broad approaches to dtruth, sketched in §4–§5, as in some sense attempting to deal with a choice between (expressive) 'completeness' and (negation) 'consistency'.]

[7]Throughout, I will restrict 'sentence' to declarative (and, for simplicity, context-independent) sentences of a given language. (In some informal examples, context-dependence creeps in, but I trust that no confusion will arise.) Unless otherwise noted, I will leave it implicit that, e.g., 'every sentence' is restricted to the sentences of the given language.

Given such a characterization—one that purports to be truly expressible in our language and likewise exhaustive—the non-triviality of our truth predicate (and, hence, our language, in general) is immediately in question in the face of Liars—the ticked sentence or the like.

The classical picture, of course, is just a special case of the Liar phenomenon. At bottom, there is a tension between the apparent non-triviality of our truth predicate and our language's apparent capacity to achieve (true) exhaustive characterization:

EC. *Exhaustive Characterization*: Every sentence is either True, False, or *Other*.

Here, 'Other' is a stand-in for the remaining 'semantically significant predicates'.[8] For present purposes, one can focus on the 'problem cases' and think of the semantically significant predicates as those that are invoked to classify such cases (e.g., Liars and the like). For example, if one wishes to classify all Liars (etc.) as *defective in some sense or other*, then 'defective in some sense or other' is semantically significant and thereby stands among one's *Others* in EC.[9]

The Liar paradox makes it difficult to see how we can have both EC and a non-trivial—let alone (negation-) consistent—truth predicate. As above, I will (until §7) focus on dtruth (or 'dtrue'), for which intersubstitutivity holds. (See §2.) The question is: How can we, as we appear to, have a non-trivial dtruth predicate (in our language, for our language) and also achieve exhaustive characterization?

In focusing on dtruth, there are two chief approaches: *paracomplete* and *paraconsistent*.[10] These are the approaches that I will discuss until deviating from dtruth at §7–§8.

4 Paracomplete

Typical Liars appear to be sentences that are equivalent to their own negations. This is especially the case with dtruth, where $\mathsf{dT}\langle A\rangle$ and A are intersubstitutable for *all* A, and hence for Liar-like A. The fundamental intersubstitutivity of $\mathsf{dT}\langle A\rangle$ and A is formally modeled by requiring sameness of semantic value: $\nu(\mathsf{dT}\langle A\rangle) = \nu(A)$ for all A.[11] Where A is a Liar sentence, the requirement is that $\nu(A) =$

[8] I will assume throughout that '(d-)true' and '(d-)false' are among our 'semantically significant predicates'. Of course, since I will also assume (throughout) that falsity is truth of negation (i.e., that A is false exactly if $\neg A$ is true), one might better put EC just as $\mathsf{True}\langle A\rangle \vee \mathsf{Other}\langle A\rangle$, with 'Other' as a stand-in, as above. But I will set this aside.

[9] I admit that the notion of 'semantically significant predicate' is not precise, but I think that the target, intuitive sense is clear. Giving a precise account is an open—and, in my opinion, quite pressing—issue. Unfortunately, except for a few passing comments, I am forced to here leave the notion fairly imprecise, trusting (hoping) that the target idea is clear enough to be useful.

[10] The terminology of 'paracomplete', explained in §4, is from Achille Varzi [1999], also used by Dominic Hyde [1997]. (The related terminology of 'gaps' and 'gluts' was introduced by Kit Fine [1975].)

[11] One can think of *intersubstitutivity* as demanding that A and B entail each other if B differs from A only in containing an occurrence of $\mathsf{dT}\langle C\rangle$ where A contains C. (Here, we assume transparent contexts.)

$\nu(\neg A)$, which is impossible in classical semantics. Paracomplete semantics afford more options.

Paracomplete accounts are the dominant approaches towards dtruth. 'Para' in 'paracomplete' comes from the Greek for *beyond* (or, perhaps, *beside*). The classical picture is one according to which every instance of LEM is dtrue, a picture according to which every A or its negation is dtrue. Paracomplete accounts reject LEM—and see it as the main principle that paradoxes call into question.[12]

While I will not discuss the issue of *rejection* or *acceptance* in this essay, it is worth noting that in paracomplete (and, indeed, paraconsistent) frameworks, 'reject A' is not usually equivalent to 'accept $\neg A$', at least if, as is the case here, we're dealing with dtruth. If one is rejecting $A \vee \neg A$ (for some A), then presumably one is thereby—or ought, rationally, thereby be—rejecting both A and $\neg A$. So, rejecting A is not to be understood as accepting that A is not dtrue, which would be equivalent to accepting $\neg A$. On all of this, paracomplete theorists can (and, presumably, will) agree.[13]

4.1 Kripke, Martin–Woodruff

The work of Kripke [1975] and Martin–Woodruff [1975] was ground-breaking, especially in a climate in which 'Tarskian approaches' were the norm. (See §7.1 for a basic sketch of Tarski's approach.) I will focus on (the non-classical) interpretation of Kripke's (least fixed point) account.[14] My aim is to give the basic philosophical picture, a sketch of the formal model, and a few comments on apparent virtues and inadequacies of the account. As throughout, I will focus on the semantic picture, tying my remarks to guiding projects NTP and ECP.

Philosophical picture

Recall the dtruth-metaphor according to which 'dtrue' is introduced for purposes of generalization. 'Prior' to introducing the device, we spoke only the 'dtrue'-free fragment. (Similarly for other semantic notions/devices, e.g., 'denotes', 'satisfies', 'true of', etc.) For simplicity, let us assume that the given 'semantic-free' fragment (hence, 'dtrue'-free fragment) is such that LEM holds.[15] Letting $\mathcal{L}_0$ be our 'semantic-free fragment', we suppose that $A \vee \neg A$ is dtrue for all A in $\mathcal{L}_0$.[16] In-

[12] Paracomplete accounts are often called 'partiality accounts' or 'partial-predicate accounts'. See [McGee, 1991], [Reinhardt, 1986], and [Soames, 1999].

[13] This will call for a modification of classical probability theory, at least if 'rejection' (similarly, acceptance) is to be understood in terms of probability (e.g., degrees of acceptance being cashed out in terms of probabilities). I will skip discussion here (though see §4.2 for a brief footnote).

[14] There are various interpretation-issues that surround Kripke's seminal work, both philosophical and logical. See §4.1 for some of the issues.

[15] This assumption sets aside the issue of vagueness (and related sorites puzzles). I am setting this aside only for simplicity. The issue of vagueness—or, as some say, 'indeterminacy', in general—is quite relevant to some paracomplete approaches to dtruth. See [Field, 2003b], [McGee, 1991], [Soames, 1999].

[16] This assumption is not essential to Kripke's account; however, it makes the basic picture much easier to see.

deed, we may suppose that classical semantics—and logic, generally—is entirely appropriate for the fragment $\mathcal{L}_0$.

But now we want our generalization-device. How do we want this to work? As above, we want $\mathsf{dT}\langle A\rangle$ and A to be intersubstitutable for *all* A. The trouble, of course, is that once 'is dtrue' is introduced into the language, various unintended—and, given the role of the device, paradoxical—sentences emerge (e.g., the ticked sentence above).[17]

The paracomplete idea, of which Kripke's is the best known, is (in effect) to allow some instances of $A \vee \neg A$ to fail.[18] In particular, if A itself fails to 'ground out' in $\mathcal{L}_0$, fails to 'find a value' by being ultimately equivalent to a sentence in $\mathcal{L}_0$, then the A-instance of LEM should fail.[19]

Kripke illustrated the idea in terms of a learning or teaching process. The guiding principle is that $\mathsf{dT}\langle A\rangle$ is to be asserted exactly when A is to be asserted. Consider an $\mathcal{L}_0$-sentence that you're prepared to assert—say, '$1 + 1 = 2$' or 'Max is a cat' or whatever. Heeding the guiding principle, you may then assert that '$1 + 1 = 2$' and 'Max is a cat' are dtrue. In turn, since you are now prepared to assert

1. 'Max is a cat' is dtrue

the guiding principle instructs that you may also assert

2. ' 'Max is a cat' is dtrue ' is dtrue.

And so on. More generally, your learning can be seen as a process of achieving further and further dtruth-attributions to sentences that 'ground out' in $\mathcal{L}_0$. (Similarly for dfalsity, which is just dtruth of negation.) Eventually, your competence reflects precisely the defining intersubstitutivity—and transparency—of dtruth: that $\mathsf{dT}\langle A\rangle$ and A are intersubstitutable for *all* A of the language.

But your competence also reflects something else: namely, the failure to assert either A or $\neg A$, for some A in the language. To see the point, think of the above process of 'further and further dtruth-attributions' as a process of writing two (very, very big) books—one, *The Truth*, the other *The False*. Think of each stage in the process as completing a 'chapter', with chapter zero of each book being empty—this indicating that *at the beginning* nothing is explicitly recorded as dtrue (or, derivately, dfalse).[20]

Concentrate just on the process of recording *atomics* in *The True*. When you were first learning, you scanned $\mathcal{L}_0$ (semantic-free fragment) for the dtrue (atomic)

[17] With respect to formal languages, the inevitability of such sentences is enshrined in Gödel's so-called *diagonal lemma*. (Even though the result is itself quite significant, it is standardly called a *lemma* because of its role in establishing Gödel–Tarski indefinability theorems. For user-friendly discussion of the limitative results, and for primary sources, see [Smullyan, 1992]. For a general discussion of diagonalization, see [Jacquette, 2004].)

[18] NB: The sense in which instances of $A \vee \neg A$ 'fail' is modeled by such instances being undesignated (in the formal model). (See §4.1.) How, if at all, such 'failure' is expressed *in* the given language concerns ECP. (See 4.1.)

[19] This is the so-called *least fixed point* picture.

[20] Publishers would probably delete chapter zero, but it's worth keeping it for present purposes.

sentences, the sentences you were prepared to assert. Chapter one of *The Truth* comprises the results of your search—sentences such as 'Max is a cat' and the like. In other words, letting '$I(t)$' abbreviate *the denotation of* t, chapter one of *The Truth* contains all of those atomics $A(t)$ such that $I(t)$ exemplifies A, a 'fact' that *would've been* recorded in chapter zero *had* chapter zero recorded the dtrue semantic-free sentences. (For simplicity, if $A(t)$ is an $\mathcal{L}_0$-atomic such that $I(t)$ exemplifies A, then we'll say that $I(t)$ exemplifies A *according to chapter zero*. In the case of 'Max is a cat', chapter zero has it that Max exemplifies cathood, even though neither 'Max is a cat' nor anything else appears in chapter zero.)

In the other book, *The False*, chapter zero is similarly empty; however, like chapter zero of *The Truth*, the sentences that *would* go into *The False*'s chapter zero are those (atomic) $\mathcal{L}_0$-sentences that, according to the world (as it were), are dfalse—e.g., '$1 + 1 = 3$', 'Max is a dog', or the like.[21] If $A(t)$ is a dfalse $\mathcal{L}_0$-atomic, we'll say that *according to chapter zero*, $I(t)$ exemplifies $\neg A$ (even though, as above, chapter zero explicitly records nothing at all). In turn, chapter one of *The False* contains all of those atomics $A(t)$ such that, according to chapter zero, $I(t)$ exemplifies $\neg A$ (i.e., the $\mathcal{L}_0$-atomics that are dfalse, even though you wouldn't say as much at this stage).

And now the writing (of atomics) continues: chapter two of *The Truth* comprises 'first-degree dtruth-attributions' and atomics $A(t)$ such that, as above, $I(t)$ exemplifies A according to chapter *one*, sentences like (1) and 'Max is a cat'. In turn, chapter three of *The Truth* comprises 'second-degree' attributions, such as (2), and atomics $A(t)$ such that (as it were) t is A according to chapter *two*. And so on, and similarly for *The False*. In general, your writing-project exhibits a pattern. Where $I_i(\mathsf{dT})$ is chapter i of *The Truth*, the pattern runs thus:

$$I_{i+1}(\mathsf{dT}) = I_i(\mathsf{dT}) \cup \{A(t) : A(t) \text{ is an atomic and } I(t) \text{ exemplifies } A \text{ according to } I_i(\mathsf{dT})\}$$

Let $\mathcal{S}$ comprise all sentences of the language. With respect to *The False* book, the pattern of your writing (with respect to atomics) looks thus:

$$I_{i+1}(\mathsf{dF}) = I_i(\mathsf{dF}) \cup \{A(t) : A(t) \text{ is an atomic and } I(t) \notin \mathcal{S} \text{ or } I(t) \text{ exemplifies } \neg A \text{ according to } I_i(\mathsf{dT})\}$$

So goes the basic process for *atomics*. But what about compound (molecular) sentences? The details are sketched below (see §4.1), but for now the basic idea is as follows (here skipping the relativizing to chapters). With respect to negations, $\neg A$ goes into *The True* just when A goes into *The False*. (Otherwise, neither A nor $\neg A$ finds a place in either book.) With respect to *conjunctions*, $A \wedge B$ goes into *The False* if either A or B goes into *The False*, and it goes into *The True* just if both A and B go into *The True*. (Otherwise, $A \wedge B$ finds a place in neither

[21] For convenience, we'll also put non-sentences into *The False*. Putting non-sentences into *The False* is not essential to Kripke's construction, but it makes things easier. Obviously, one can't *write* a cat but, for present purposes, one can think of *The False* as a special book that comes equipped with attached nets (wherein non-sentences go), a net for each chapter.

book.) The case of *disjunctions* is dual, and the quantifiers may be treated as 'generalized conjunction' (universal) and 'generalized disjunction' (existential).[22]

Does every sentence eventually find a place in one book or other? No. Consider an atomic sentence L, like the ticked sentence in §3, equivalent to $\neg \mathsf{dT}\langle L\rangle$. In order to get L into *The True* book, there'd have to be some chapter in which it appears. L doesn't appear in chapter zero, since nothing does. Moreover, L doesn't exemplify anything 'according to chapter zero', since chapter zero concerns only the $\mathcal{L}_0$-sentences (and L isn't one of those). What about chapter one? In order for L to appear in chapter one, L would have to be in chapter zero or be such that L exemplifies $\neg \mathsf{dT}(x)$ according to chapter zero. But for reasons just given, L satisfies neither disjunct, and so doesn't appear in chapter one. The same is evident for chapter two, chapter three, and so on. Moreover, the same reasoning indicates that L doesn't appear in *The False* book.

In general, Liar-like sentences such as the ticked sentence in §3 will find a place in one of our books only if it finds a place in one of the chapters $I_i(\mathsf{dT})$ or $I_i(\mathsf{dF})$. But the ticked sentence will find a place in $I_i(\mathsf{dT})$ or $I_i(\mathsf{dF})$ only if it finds a place in $I_{i-1}(\mathsf{dT})$ or $I_{i-1}(\mathsf{dF})$. But, again, the ticked sentence will find a place in $I_{i-1}(\mathsf{dT})$ or $I_{i-1}(\mathsf{dF})$ only if it finds a place in $I_{i-2}(\mathsf{dT})$ or $I_{i-2}(\mathsf{dF})$. And so on. But, then, since $I_0(\mathsf{dT})$ and $I_0(\mathsf{dF})$ are both empty, and since—by our stipulation—something exemplifies a property according to $I_0(\mathsf{dT})$ only if the property is a non-semantic one (the predicate is in $\mathcal{L}_0$), the ticked sentence (or the like) fails to find a place in either book. Such a sentence, according to Kripke, is not only *ungrounded*, since it finds a place in neither book, but also *paradoxical*—it *couldn't* find a place in either book.[23]

So goes the basic philosophical picture. What was wanted was an account of how, despite the existence of Liars, we could have a dtruth predicate in the language—and do so without triviality (or, in Kripke's case, inconsistency). The foregoing picture suggests an answer, at least if we eventually have a chapter $I_i(\mathsf{dT})$ such that $\mathsf{dT}\langle A\rangle$ is in $I_i(\mathsf{dT})$ if and only if A is in $I_i(\mathsf{dT})$, and similarly a chapter for *The False*. What Kripke (and, independently, Martin–Woodruff) showed is that, provided our 'writing process' follows the right sort of scheme (in effect, a logic weaker than classical), our books will contain such target chapters, and in that respect our language can enjoy a (non-trivial, indeed consistent) dtruth predicate. Making the philosophical picture more precise is the job of formal, philosophical modeling, to which I now briefly turn.

Formal model

The main results of [Kripke, 1975] and [Martin and Woodruff, 1975] are more general than I will discuss here. As throughout, my emphasis attempts to be

[22]This approach to compound sentences reflects the so-called *Strong Kleene* scheme, which is given in §4.1.

[23]The force of *couldn't* here is made precise by the full semantics, but for present purposes one can think of *couldn't* along the lines of *on pain of (negation-) inconsistency* or, for that matter, *on pain of being in both books* (something impossible, on the current framework).

philosophical, with an aim towards 'real truth', with answering NTP and ECP with respect to our 'real language'.

Of course, even though our aim is 'real language', we must none the less abstract a bit from the mess. The aim of formal accounts of truth, at least those concerned with 'real truth' in natural languages (or the very language 'we' speak), is not to give an account of truth, but rather truth-in-$\mathcal{L}$, for some formal 'model-language' $\mathcal{L}$. The relevance of such an account is that 'real truth' is supposed to be 'similar enough' to $\mathcal{L}$-truth, at least in relevant respects, to gain answers to NTP and, perhaps, ECP.

For present purposes, I focus on what is known as Kripke's 'least fixed point' model (with empty ground model). I leave proofs to cited works (all of which are readily available), and try to say just enough to see how the formal picture goes.

Following standard practice, we can think of an *interpreted language* $\mathcal{L}$ as a triple $\langle \mathtt{L}, \mathcal{M}, \sigma\rangle$, where $\mathtt{L}$ is the syntax (the relevant syntactical information), $\mathcal{M}$ is an 'interpretation' or 'model' that provides interpretations to the non-logical constants (names, function-symbols, predicates), and σ is a 'semantic scheme' or 'valuation scheme' that, in effect, provides interpretations—semantic values—to compound sentences.[24]

Consider, for example, familiar classical languages, where the set $\mathcal{V}$ of 'semantic values' is $\{1, 0\}$. In classical languages, $\mathcal{M} = \langle \mathcal{D}, I\rangle$, with $\mathcal{D}$ our (non-empty) domain and I an 'interpretation-function' that assigns to each name an element of $\mathcal{D}$ (the denotation of the name), assigns to each n-ary function-symbol an element of $\mathcal{D}^n \longrightarrow \mathcal{D}$, that is, an n-ary function from $\mathcal{D}^n$ into $\mathcal{D}$, and assigns to each n-ary predicate an element of $\mathcal{D}^n \longrightarrow \mathcal{V}$, a function—sometimes thought of as the *intension* of the predicate—taking n-tuples of $\mathcal{D}$ and yielding a 'semantic value' (a 'truth value'). The *extension* of an n-ary predicate F (intuitively, the set of things of which F is true) contains all n-tuples $\langle a_1, \ldots, a_n\rangle$ of $\mathcal{D}$ such that $I(F)(\langle a_1, \ldots, a_n\rangle) = 1$. The classical valuation scheme τ (for Tarski) is the familiar one according to which a negation is true (in a given model) exactly when its negatum is false (in the given model), a disjunction is true (in a model) iff one of the disjuncts is true (in the model), and existential sentences are treated as generalized disjunctions.[25]

Classical languages (with suitably resourceful $\mathtt{L}$) cannot have their own dtruth predicate, which is the upshot of Tarski's indefinability theorem. Paracomplete languages reject the 'exhaustive' feature implicit in classical languages: namely, that a sentence or its negation is true, for *all* sentences.

The standard way of formalizing paracomplete languages expands the interpretation of predicates. Recall that in your 'writing process' some sentences (e.g.,

[24] For present purposes, a semantic scheme or valuation scheme σ is simply some general definition of *truth (falsity) in a model.* For a more detailed discussion, see [Gupta and Belnap, 1993].

[25] I assume familiarity with the basic classical picture, including 'true in $\mathcal{L}$' and so on. To make things easier, I will sometimes assume that we've moved to models in which everything in the domain has a name, and otherwise I'll assume familiarity with standard accounts of 'satisfies $A(x)$ in $\mathcal{L}$'.

Liars) found a place in neither book. We need to make room for such sentences, and we can expand our semantic values $\mathcal{V}$ to do so; we can let $\mathcal{V} = \{1, \frac{1}{2}, 0\}$, letting the middle value represent the status of sentences that found a place in neither book.

Generalizing (but, now, straining) the metaphor, we can think of all n-ary predicates as tied to two such 'big books', one recording the objects of which the predicate is true, the other the objects of which it is false. On this picture, the *extension* of a predicate F remains as per the classical (containing all n-tuples of which the predicate is true), but we now also acknowledge an *antiextension*, this comprising all n-tuples of which the predicate is false. This broader picture of predicates enjoys the classical picture as a special case: namely, where we stipulate that, for any predicate, the extension and antiextension are jointly exhaustive (the union of the two equals the domain) and, of course, exclusive (the intersection of the two is empty).

Concentrating on the so-called *Strong Kleene* account [1952],[26] the formal story runs as follows. We expand $\mathcal{V}$, as above, to be $\{1, \frac{1}{2}, 0\}$, and so our language $\mathcal{L}_\kappa = \langle \mathtt{L}, \mathcal{M}, \kappa \rangle$ is now a so-called three-valued language (because it uses three semantic values).[27] Our *designated values*—intuitively, the values in terms of which *validity* or *consequence* is defined—are a subset of our semantic values; in the Strong Kleene case, there is exactly one designated element, namely 1.

A (Strong Kleene) model $\mathcal{M} = \langle \mathcal{D}, I \rangle$ is much as before, with I doing exactly what it did in the classical case except that I now assigns to n-ary predicates elements of $D^n \longrightarrow \{1, \frac{1}{2}, 0\}$, since $\mathcal{V} = \{1, \frac{1}{2}, 0\}$. Accordingly, the 'intensions' of our paracomplete (Strong Kleene) predicates have three options: 1, $\frac{1}{2}$, and 0. What about *extensions*? As above, we want to treat predicates not just in terms of extensions (as in the classical languages) but also antiextensions. The *extension* of an n-ary predicate F, just as before, comprises all n-tuples $\langle a_1, \ldots, a_n \rangle$ of $\mathcal{D}$ such that $I(F)(\langle a_1, \ldots, a_n \rangle) = 1$. (Again, intuitively, this remains the set of objects of which F is true.) The *antiextension*, in turn, comprises all n-tuples $\langle a_1, \ldots, a_n \rangle$ of $\mathcal{D}$ such that $I(F)(\langle a_1, \ldots, a_n \rangle) = 0$. (Again, intuitively, this is the set of objects of which F is false.) Of course, as intended, an interpretation might fail to put x in either the extension or antiextension of F. In that case, we say (in our 'metalanguage') that, relative to the model, F is *undefined* for x.[28]

[26]This is one of the paracomplete languages for which Kripke proved his definability result. Martin–Woodruff proved a special case of Kripke's general 'fixed point' result, namely, the case for so-called 'maximal fixed points' of the *Weak Kleene* scheme, or *weak Kleene languages*.

[27]Kripke [1975] made much of emphasizing that 'the third value' is not to be understood as *a third truth value* or anything else other than 'undefined' (along the lines of Kleene's original work [1952]). I will not make much of this here, although what to make of semantic values that appear in one's formal account is an important, philosophical issue, one that I'll briefly touch on in §10. (Note that if one wants to avoid a three-valued language, one can let $\mathcal{V} = \{1, 0\}$ and proceed to construct a Kleene-language by using *partial functions* (hence, the standard terminology 'partial predicates') for interpretations. I think that this is ultimately merely terminological, but I won't dwell on the matter here.

[28]A common way of speaking is to say that, for example, $F(t)$ is 'gappy' with respect to $I(t)$. This terminology is appropriate if one is clear on the relation between one's formal model and the

Letting $\mathcal{F}^+$ and $\mathcal{F}^-$ be the extension and antiextension of F, respectively, it is easy to see that, as noted above, classical languages are a special case of (Strong Kleene) paracomplete languages. Paracomplete languages typically eschew inconsistency,[29] and so typically demand that $\mathcal{F}^+ \cap \mathcal{F}^- = \emptyset$, in other words, that nothing is in both the extension and antiextension of any predicate. In this way, paracomplete languages (typically) agree with classical languages. The difference, of course, is that paracomplete languages do *not* demand that $\mathcal{F}^+ \cup \mathcal{F}^- = \mathcal{D}$ for all predicates F. But paracomplete languages *allow* for such 'exhaustive constraints', and in that respect can enjoy classical languages as a special case.

To see the close relation between classical languages and Strong Kleene, notice that κ, the Strong Kleene valuation-scheme, runs as follows (here treating only $\neg$, $\vee$, and $\exists$). Where $V_{\mathcal{M}}(A)$ is the semantic value of A in $\mathcal{M}$ (and, for simplicity, letting each object in the domain name itself), and, for purposes of specifying scheme κ, treating $\mathcal{V}$ as standardly (linearly) ordered:

K1. $V_{\mathcal{M}}(\neg A) = 1 - V_{\mathcal{M}}(A)$.

K2. $V_{\mathcal{M}}(A \vee B) = \mathtt{max}(V_{\mathcal{M}}(A), V_{\mathcal{M}}(B))$.

K3. $V_{\mathcal{M}}(\exists x\, A(x)) = \mathtt{max}\{V_{\mathcal{M}}(A(t/x)) : \text{ for all } t \in \mathcal{D}\}$.

The extent to which classical logic is an extension of a given paracomplete logic depends on the semantic scheme of the language.[30] Since κ, as above, is entirely in keeping with the classical scheme τ *except* for 'adding an extra possibility', it is clear that every classical interpretation is a Strong Kleene-interpretation (but not vice-versa).[31]

Let us say that an interpretation *verifies* a sentence A iff A is designated (in this case, assigned 1) on that interpretation, and that an interpretation verifies a set of sentences Σ iff it verifies every element of Σ. We define *semantic consequence* in familiar terms: A is a consequence of Σ iff every interpretation that verifies Σ also verifies A. I will use '$\Vdash_\kappa$' for the Strong Kleene consequence relation, so understood.[32]

Let us say that a sentence A is logically true in $\mathcal{L}_\kappa$ exactly if $\emptyset \Vdash_\kappa A$, that is, iff A is designated (assigned 1) in every model. A remarkable feature of $\mathcal{L}_\kappa$ is that there are no logical truths. To see this, just consider an interpretation that

target notions that the model is intended to serve (in one respect or other), but the terminology can also be confusing. I will return to this point below (see §4.2 and §10).

[29] But see §5.2.

[30] Here, perhaps not altogether appropriately, I am privileging model theory over proof theory, thinking of 'logic' as the semantic consequence relation that falls out of the semantics. This is in keeping with the elementary aims of the essay, even though (admittedly) it blurs over a lot of philosophical and logical issues.

[31] Note that in classical languages, $V_{\mathcal{M}}(A) \in \{1, 0\}$ for any A, and the familiar classical clauses on negation, disjunction, etc. are then simply (K1)–(K3).

[32] '$\Vdash$' is increasingly used for *forcing* in set theoretic-logical literature, and so '$\models$' might be a better choice for semantic consequence. The trouble is that the latter is very standardly used in truth-theoretic literature to represent the *true-in-a-language* or *true-in-a-model* relation.

assigns $\frac{1}{2}$ to every atomic, in which case, as an induction will show, every sentence is assigned $\frac{1}{2}$ on that interpretation. Hence, there's some interpretation in which no sentence is designated, and hence no sentence designated on all interpretations. A fortiori, LEM fails in Strong Kleene languages.[33]

And now an answer to NTP, with respect to dtruth, becomes apparent. What we want is a model of how our language can be non-trivial (indeed, consistent) while containing both a dtruth predicate and Liar-like sentences. In large part, the answer is that our language is (in relevant respects) along Strong Kleene lines, that the logic is weaker than classical logic. Such a language, as Kripke showed, can contain its own dtruth predicate.

The construction runs (in effect) along the lines of the 'big books' picture. For simplicity, let $\mathcal{L}_\kappa$ be a *classical* (and, hence, Strong Kleene) language such that $\mathtt{L}$ (the basic syntax, etc.) is free of semantic terms but has the resources to describe its given syntax—including, among other things, having a name $\langle A \rangle$ for each sentence. (So, I assigns to each n-ary predicate an element of $\mathcal{D}^n \longrightarrow \{1, 0\}$, even though the values $\mathcal{V}$ of $\mathcal{L}_\kappa$ also contain $\frac{1}{2}$.) What we want to do is move to a richer language the syntax $\mathtt{L}^t$ of which contains $\mathsf{dT}(x)$, a unary predicate (intended to be a dtruth predicate for the enriched language). For simplicity, assume that the domain $\mathcal{D}$ of $\mathcal{L}_\kappa$ contains all sentences of $\mathtt{L}^t$.[34]

Think, briefly, about the 'big books' picture. One can think of each successive 'chapter' as a language that expands one's official record of what is dtrue (dfalse). More formally, one can think of each such 'chapter' of both books as the extension and antiextension of 'dtrue', with each such chapter expanding the interpretation of 'dtrue'. Intuitively (with slight qualifications about chapters zero), one can think of $I_{i+1}(\mathsf{dT})$ as explicitly recording 'what is true' *according to* chapter $I_i(\mathsf{dT})$. The goal, of course, is to find a 'chapter' at which we have $I_{i+1}(\mathsf{dT}) = I_i(\mathsf{dT})$, a 'fixed point' at which anything dtrue in the language is fully recorded in the given chapter—one needn't go further. Thinking of the various 'chapters' as *languages*, each with a richer interpretation of 'dtrue', one can think of the 'fixed chapter' as a language that, finally, has a dtruth predicate for itself.

Returning to the construction at hand, we have our Strong Kleene (but classical) 'ground language' $\mathcal{L}_\kappa$ that we now expand to $\mathcal{L}_\kappa^t$, the syntax of which includes that of $\mathcal{L}_\kappa$ but also has $\mathsf{dT}(x)$ (and the resulting sentences formable therefrom). We want the new language to 'expand' the ground language, and we want the former to have a model that differs from the latter only in that it assigns an interpretation to $\mathsf{dT}(x)$. For present purposes, we let I^t, the interpretation function in $\mathcal{L}_\kappa^t$, assign $(\emptyset, \emptyset)$ to $\mathsf{dT}(x)$, where $(\emptyset, \emptyset)$ is the function that assigns $\frac{1}{2}$ to each element of $\mathcal{D}^t$. (Hence, the extension and antiextension of $\mathsf{dT}(x)$ in $\mathcal{L}_\kappa^t$ are both empty.) This is the formal analogue of 'chapter zero'.

[33] This is not to say, of course, that one can't have a Strong Kleene—or paracomplete, in general—language some proper fragment of which is such that $A \vee \neg A$ holds for all A in the fragment. (One might, e.g., stipulate that arithmetic is such that $A \vee \neg A$ holds.)

[34] This is usually put (more precisely) as that the domain contains the Gödel-codes of all such sentences, but for present purposes I will skip over the mathematical details.

The crucial question, of course, concerns *further expansion*. How do we expand the interpretation of $\mathsf{dT}(x)$? How do we move to 'other chapters'? How, in short, do we eventually reach a 'chapter' or language in which we have a dtruth predicate for the whole given language? This is the role of Kripke's 'jump operator'. What we want, of course, are 'increasingly informative' interpretations $(\mathcal{T}_i^+, \mathcal{T}_i^-)$ of $\mathsf{dT}(x)$, but interpretations that not only 'expand' the previous interpretations but also *preserve* what has already been interpreted. If A is dtrue according to chapter i, then we want as much preserved: that A remain dtrue according to chapter $i+1$. This is the role of the 'jump operator', a role that is achievable given the so-called *monotonicity* of Strong Kleene valuation scheme κ.[35] The role of the jump operator is to eventually 'jump' through successive interpretations (chapters, languages) $I_i(\mathsf{dT})$ and land on one that serves the role of dtruth—serves as an interpretation of 'dtrue'. As above, letting $I_i(\mathsf{dT})$ be a function $(\mathcal{T}_i^+, \mathcal{T}_i^-)$ yielding 'both chapters i', the goal is to eventually 'jump' upon an interpretation $(\mathcal{T}_i^+, \mathcal{T}_i^-)$ such that $(\mathcal{T}_i^+, \mathcal{T}_i^-) = (\mathcal{T}_{i+1}^+, \mathcal{T}_{i+1}^-)$.

Focusing on the 'least such point' in the Strong Kleene setting, Kripke's construction proceeds as above. We begin at stage 0 at which dT is interpreted as $(\emptyset, \emptyset)$, and we define a 'jump operator' on such interpretations:[36] dT is interpreted as $(\mathcal{T}_{i+1}^+, \mathcal{T}_{i+1}^-)$ at stage $i+1$ if interpreted as $(\mathcal{T}_i^+, \mathcal{T}_i^-)$ at the preceding stage i, where, note well, $\mathcal{T}_{i+1}^+$ comprises the sentences that are true (designated) at the preceding stage (chapter, language) i, and $\mathcal{T}_{i+1}^-$ the false sentences (and, for simplicity, non-sentences) at i. Accordingly, we define the 'jump operator' J_κ thus:[37]

$$J_\kappa(\mathcal{T}_i^+, \mathcal{T}_i^-) = (\mathcal{T}_{i+1}^+, \mathcal{T}_{i+1}^-)$$

The jump operator yields a sequence of richer and richer interpretations that 'preserve prior information' (given monotonicity), a process that can be extended into the transfinite to yield a sequence

$$(\mathcal{T}_0^+, \mathcal{T}_0^-), (\mathcal{T}_1^+, \mathcal{T}_1^-), \ldots, (\mathcal{T}_\alpha^+, \mathcal{T}_\alpha^-), \ldots$$

defined (via transfinite recursion) thus:[38]

[35] Monotonicity is the crucial ingredient in Kripke's (similarly, Martin–Woodruff's) general result. Let $\mathcal{M}$ and $\mathcal{M}'$ be paracomplete (partial) models for (uninterpreted) $\mathtt{L}$. Let $\mathcal{F}_M^+$ be the extension of F in $\mathcal{M}$, and similarly $\mathcal{F}_{M'}^+$ for $\mathcal{M}'$. (Similarly for antiextension.) Then $\mathcal{M}'$ *extends* $\mathcal{M}$ iff the models have the same domain, agree on interpretations of names and function signs, and $\mathcal{F}_M^+ \subseteq \mathcal{F}_{M'}^+$ and $\mathcal{F}_M^- \subseteq \mathcal{F}_{M'}^-$ for all predicates F that M and M' interpret. (In other words, $\mathcal{M}'$ doesn't change $\mathcal{M}$'s interpretation; it simply interprets whatever, if anything, $\mathcal{M}$ left uninterpreted.) MONOTONICITY PROPERTY: A semantic (valuation) scheme σ is *monotone* iff for any A that is interpreted by both models, A's being designated in $\mathcal{M}$ implies its being designated in $\mathcal{M}'$ whenever $\mathcal{M}'$ extends $\mathcal{M}$. So, the montonicity property of a scheme ensures that it 'preserves truth (falsity)' of 'prior interpretations' in the desired fashion.

[36] So, our operator operates on the set of all (admissible) functions from $\mathcal{D}$ into $\{1, \frac{1}{2}, 0\}$, where $\mathcal{D}$ is in our 'ground language', as above.

[37] Note that Kripke's definition applies to *any* monotone scheme σ. I relativize the operator to κ just to remind that we here focusing on the Strong Kleene case.

[38] If the reader isn't familiar with transfinite recursion, just note that it's much like ordinary

Jb. Base. $(\mathcal{T}_0^+, \mathcal{T}_0^-) = (\emptyset, \emptyset)$.

Js. Successor. $(\mathcal{T}_{\alpha+1}^+, \mathcal{T}_{\alpha+1}^-) = J_\kappa((\mathcal{T}_\alpha^+, \mathcal{T}_\alpha^-))$.

Jl. Limit. For limit stages, we collect up by unionising the prior stages:

$$(\mathcal{T}_\lambda^+, \mathcal{T}_\lambda^-) = \left(\bigcup_{\epsilon<\lambda} \mathcal{T}_\epsilon^+, \bigcup_{\epsilon<\lambda} \mathcal{T}_\epsilon^-\right)$$

What Kripke showed—for *any* monotone scheme, and a fortiori for Strong Kleene—is that the transfinite sequence reaches a stage at which the desired dtruth predicate is found, a 'fixed point' of the jump operator such that we obtain

$$(\mathcal{T}_\alpha^+, \mathcal{T}_\alpha^-) = (\mathcal{T}_{\alpha+1}^+, \mathcal{T}_{\alpha+1}^-) = J_\kappa((\mathcal{T}_\alpha^+, \mathcal{T}_\alpha^-))$$

The upshot is that 'chapter α' or 'language α' is such that $\mathcal{T}_\alpha^+$ and $\mathcal{T}_\alpha^-$ comprise all of the dtrue (respectively, dfalse) sentences of $\mathcal{L}_\kappa^\alpha$, the language at α, which is to say that $\mathcal{L}_\kappa^\alpha$ contains its own dtruth predicate.

The *proof* of Kripke's result is left to cited work.[39] For present purposes, I move on to a few comments about the given paracomplete account.

Comments

In this section I briefly mention a few issues concerning the philosophical application of Kripke's paracomplete account, concentrating on the unifying projects, NTP and ECP. I leave much of the discussion to cited sources, especially with respect to what I call 'interpretation issues'.

Interpretation issues

Three salient interpetation-related issues have emerged with respect to Kripke's proposal(s).[40] I simply mention the issues here, pointing to cited works for further discussion.[41]

---

recursive definitions except for requiring an extra clause for so-called 'limit ordinals'. Here, α and ϵ are ordinals, and λ a 'limit ordinal'. (One can find a discussion of transfinite recursion in most standard set theory books or metatheory textbooks. See too [McGee, 1991], [Smullyan, 1993].)

[39]Kripke's own proof is elegant, bringing in mathematically important and interesting results of recursion theory and inductive definitions. Kripke's proof is also perhaps more philosophically informative than a popular algebraic proof, especially with respect to the least fixed point (on which we've focused here). Still, if one simply wants *a* proof of the given result (e.g., existence of least fixed point), a straightforward algebraic proof is available, due to Visser [2004] and Fitting [1986], and discussed in a general, user-friendly fashion by Gupta–Belnap [1993].

[40]Despite my talk of *Kripke's proposal* or the like, I should make the traditional note that Kripke himself abstained from endorsing any of the particular accounts in his [Kripke, 1975]. So noted.

[41]I should perhaps emphasize that, by my lights, the 'interpretation issues' are important and, were it not for space considerations, I would devote much more space to them.

Ik1. Which fixed point?

Ik2. Supervenience or Transparency?

Ik3. Classical or non-classical theory?

I will (very) briefly discuss each issue in turn.

(Ik1): Take any monotonic semantic (valuation) scheme σ.[42] Kripke showed that the (suitably defined) 'jump operator' over σ-interpretations will have a fixed point that can serve as an interpretation of a dtruth predicate (for the given language). But while we've narrowly focused on one particular interpretation (the least fixed point), there are in fact many fixed points (as Kripke noted), and there's some controversy about which of the many fixed points best model the dtruth predicate. For discussion of the issue see [Gupta and Belnap, 1993].

(Ik2): This issue is related to (Ik1). Michael Kremer [1988] argues that the so-called 'supervenience' and 'transparency' ideas about dtruth are in conflict in Kripke's proposal. The former is the idea that once the non-semantic 'facts' are fixed, then so too is the interpretation of 'dtrue'. The latter idea is as indicated throughout: that 'dtrue' is entirely transparent, that it, unlike the former idea, dictates no particular interpretation other than one that affords its essential intersubstitutivity. A related issue—concerning the philosophical significance of fixed point constructions, generally—is discussed by Philip Kremer [2000].

(Ik3): This is a slightly more technical issue. In §4.1–§4.1 I focused entirely on κ-based languages, the Strong Kleene, least fixed point proposal. Moreover, I have focused entirely on a non-classical reading of that proposal—one for which the resulting logic is non-classical (and, indeed, Strong Kleene). But a related, classical reading is also well-known, a reading (and resulting classical truth theory) proposed by Sol Feferman [1984]. This reading is standardly called 'KF' (for Kripke–Feferman). While the proposal is interesting, it is not an account of *dtruth*, since it gives up the essential intersubstitutivity of $\mathsf{dT}\langle A\rangle$ and A.[43] See [Reinhardt, 1986] and [Maudlin, 2004] for further discussion and related proposals.

Non-triviality project

Recall that NTP, the non-triviality project, is to explain how we can enjoy a non-trivial language that has a dtruth predicate and Liar-like sentences. In Kripke's case, the project is to show how we can enjoy a *consistent* dtruth predicate despite Liars. In general, one aims to answer NTP by constructing an artificial, formal language—the *model language*—that contains its own dtruth predicate, and then

[42]Kripke [1975] explicitly discusses two other well-known schemes, namely, the Weak Kleene scheme (which is Bochvar's 'internal' scheme) and van Fraassen's *supervaluational* scheme. (Martin–Woodruff [1975] proved the existence of the 'maximal fixed point' wrt the Weak Kleene scheme.)

[43]That it gives up on dtruth is plain from the fact that Feferman was after a consistent, classical theory, something with which dtruth is incompatible, at least given usual syntactic resources. (One conspicuous example of the departure from dtruth is in KF's commitment to instances of $A \wedge \neg\mathsf{Tr}\langle A\rangle$, something that would be inconsistent if 'Tr' were read as dtruth.)

claims that, at least in relevant respects, 'real dtruth' is modeled by truth-in-the-model language. The point of the model language is to serve as a clear, albeit idealized, model that explains how dtruth (in our real language) can achieve consistency (or, more broadly, non-triviality).[44]

Some standardly object that Kripke's account doesn't answer NTP, the reason being that certain notions used in the 'metalanguage' (the language in which one constructs one's 'model language') are not expressible in the object- or model language. I will return to this issue—what is often dubbed *revenge*—in §10. For now, I will simply sketch an example of a common charge against Kripke's (and many similar) accounts.

Concentrate on the canvassed case, namely, the Strong Kleene case. Let $\mathcal{L}_\kappa$ be such a (fixed point) language constructed via the κ-scheme.[45] In constructing $\mathcal{L}_\kappa$, we use—in our metalanguage—classical set theory, and we define *truth-in-*$\mathcal{L}_\kappa$ (and similarly, $\mathcal{L}_\kappa$-falsity). Moreover, we can prove—in our metalanguage—that, despite paradoxical sentences, a sentence $\mathsf{dT}\langle A\rangle$ is $\mathcal{L}_\kappa$-true exactly if A is $\mathcal{L}_\kappa$-true.

The common charge is that $\mathcal{L}_\kappa$, so understood, is not an adequate account of dtruth (dfalsity) itself: it fails to illuminate how dtruth itself achieves consistency. The charge is that $\mathcal{L}_\kappa$-truth achieves its consistency in virtue of $\mathcal{L}_\kappa$'s expressive poverty: $\mathcal{L}_\kappa$ cannot, on pain of inconsistency, express certain notions *that our 'real language' can express.* Example: Suppose that $\mathcal{L}_\kappa$ contains a formula $A(x)$ that defines $\{B : B \text{ is not } \mathcal{L}_\kappa\text{-true}\}$. And now, where λ says $A(\langle\lambda\rangle)$,[46] we can immediately prove—in the metalanguage—that λ is $\mathcal{L}_\kappa$-true iff $A(\langle\lambda\rangle)$ is $\mathcal{L}_\kappa$-true iff λ is not $\mathcal{L}_\kappa$-true. Because—and only because—we have it *in our classical metalanguage* that λ is $\mathcal{L}_\kappa$-true or not, we thereby have a contradiction: that λ is both $\mathcal{L}_\kappa$-true and not. But since we have it that $\mathcal{L}_\kappa$ is consistent (given consistency of classical set theory in which $\mathcal{L}_\kappa$ is constructed), we conclude that $\mathcal{L}_\kappa$ cannot express 'is not $\mathcal{L}_\kappa$-true'.

The common charge, then, amounts to this: that the Kripkean model fails to be 'enough like our real language' to explain at least one of the target phenomena, namely, dtruth's consistency. Our metalanguage is part of our 'real language', and we can define $\{B : B \text{ is not } \mathcal{L}_\kappa\text{-true}\}$ in our metalanguage. As the Kripkean language cannot similarly define $\{B : B \text{ is not } \mathcal{L}_\kappa\text{-true}\}$, the Kripkean model thereby fails to illuminate dtruth's consistency. In short, the proposed model fails to adequately answer NTP.

Despite the popularity of this sort of charge, I think that it is confused. I will return to this in §10, but the basic confusion arises from conflating the model-relative notions (e.g., truth-in-the-model language), which are defined in a classical metalanguage, and the target notions in the 'real language'. Since classical logic is an extension of Strong Kleene—semantically put, any classical interpretation is a Strong Kleene interpretation—there's no obvious reason why one can't 'stand

[44]Moreover, and importantly, such artificial 'model languages' aim to serve as *models* in a more technical sense—affording a *consistency proof.*

[45]The point applies to any of the given languages.

[46]This is the familiar sense of 'says' introduced in [Burge, 1979].

squarely within a classical fragment' of one's 'real language' and define model-relative notions that serve as heuristic tools (serve, e.g., to answer NTP).[47] If that is correct, then one can't expect 'true-in-the-model language' to be exactly like dtruth, since—on the paracomplete proposal—LEM fails for the latter but holds for the former. Indeed, the difference between the two notions is conspicuous: dtruth cannot be defined in a classical language (or fragment thereof), but 'true-in-the-model language' can be so defined (e.g., in Kripke's and others' proposals). So, even on the surface, there is reason to think that the common charge (sketched above) turns on a confusion.[48]

Put another way, the problem with such common charges is that they're either confused or unwarranted. The paracomplete theorist proposes that LEM fails for negation (and, presumably, any 'negation-like' devices) in our 'real language', the language for which NTP arises. In the Strong Kleene framework, the paracomplete theorist claims that Strong Kleene is our 'real logic', the logic governing our real language. But that's compatible with the real language having an entirely classical proper fragment the logic of which—restricted to that fragment—is classical. (Recall, as above, that classical logic is a proper extension of Strong Kleene.) For convenience (or, perhaps, other reasons), the paracomplete theorist, towards answering NTP, 'stands squarely in the classical fragment' of her real language to construct an artificial language—a model language intended to model the real language in relevant respects. The construction consists of defining various model-relative notions such as *true-in-the-model language* and, note well, *its classical complement* (viz., 'not true-in-the-model language') for which, of course, LEM holds. Now, the common charge, as sketched above, is *confused* if it conflates the model-relative notion of 'not true-in-the-model language' with, for example, 'not dtrue'. But it should also be plain that the charge is *unwarranted* if it merely assumes that there's some negation-like device NoT in the real language for which, for example, every sentence (in the real language) is either dtrue or NoT. The paracomplete theorist maintains that, perhaps in addition to other phenomena, the paradoxes teach us that there's no such device. The common charge, in turn, must come equipped with an argument that there is such a device, if it is wishes to show the inadequacy of the paracomplete account with respect to NTP. (Pointing

[47]In terms of the logics, one can think of a logic being an extension of another iff the former 'contains more validities'. More precisely, letting $\mathbb{V}_L$ and $\mathbb{V}_{L'}$ comprise the valid sentences and 'rules' (argument-validities) of two different logics L and L', we say that L' extends L exactly if $\mathbb{V}_{L'} \subseteq \mathbb{V}_L$, and *properly extends* iff $\mathbb{V}_{L'} \subset \mathbb{V}_L$. As in §4.1, it is easy to see that classical logic (properly) extends Strong Kleene. (What mightn't be as obvious is that all classically valid rules that don't turn on classically valid sentences are valid in Strong Kleene.)

[48]Hartry Field [2003b], correctly in my opinion, conjectures that some of the noted confusion, at least targeted at the Kripkean Strong Kleene proposal(s), may well arise due to a confusion between the non-classical reading, on which I've focused here, and the KF-reading noted above. On the latter reading, the proposal has it that some sentences are *neither true nor false*. Such a notion makes little sense if 'true' and 'false' are understood as *dtruth* and *dfalse*. (This, of course, is precisely not how 'true' or 'false' are understood in KF!) On the non-classical reading, in which we have dtruth, it makes little sense to talk in terms of 'neither dtrue or dfalse'. I will return to the issue of 'neither dtrue or dfalse' in the next (sub- sub-) section, and also in §4.2.

out that we use 'not true-in-the-model language' in our real language is irrelevant, since that, as above, is merely model-relative.) So, such common charges are either unwarranted or confused.[49]

Despite the confusion or lack of warrant in such common charges, there is none the less something correct in the objection, something about expressive poverty. But the issue, I think, has less to do with NTP than with ECP, to which I now (briefly) turn.

Exhaustive characterization project

ECP is the project of explaining how, if at all, we can achieve 'exhaustive characterization' in a language with its own dtruth predicate (in the language) and Liar-like sentences. While 'exhaustive characterization' remains imprecise, the intuitive import is clear: a language in which we have various 'semantically significant predicates' that may be used to *exhaustively* and correctly (semantically) categorize all sentences of the language.

One way of thinking about 'exhaustive characterization', as here intended, is as follows. Suppose that our semantic-free fragment $\mathcal{L}_0$ is exhaustively characterized *classically*, in which case we have it that, where CEC is the 'classical exhaustive characterization',

CEC. Every sentence of $\mathcal{L}_0$ is either dtrue or dfalse.

Now, what Kripke showed—and paracomplete theorists, in general, advance—is that we can retain our dtruth device (the predicate 'dtrue') if our language goes beyond the classical—opens up the 'semantically significant options' for sentences. The idea is that, while CEC may suffice for $\mathcal{L}_0$, a genuinely *exhaustive* characterization requires another category:

EC. Every sentence (in the language) is either dtrue, dfalse, or *Other*.

The question is: how shall *other* be understood?

To begin, one way that 'Other' should *not* be understood is as implying *not dtrue* or, hence, *neither dtrue nor dfalse*, at least in paracomplete accounts in which normal De Morgan principles (e.g., distribution-like principles) hold for the extensional connectives.[50] Suppose, for example, that 'other' in EC is cashed out such that $\mathsf{Other}\langle A\rangle$ implies $\neg\mathsf{dT}\langle A\rangle$, and hence—given intersubstitutivity—implies $\neg A$. One reason for introducing 'Other' into the language is to *correctly characterize* Liars—e.g., the ticked sentence in §3. Let A be sentence that says of itself (only) that it's not dtrue, that is, a sentence equivalent to $\neg\mathsf{dT}\langle A\rangle$ and,

[49] Hartry Field [2003b; 2005b] has made a lot of use—good use, in my opinion—of this basic point. But I trust that the main point is fairly clear and uncontroversial. (Where controversy might arise is with ECP, to which I turn below.)

[50] In my discussion of paracomplete (and, for that matter, paraconsistent) accounts, I focus entirely on 'normal accounts', accounts for which conjunction, disjunction, and negation interact in standard ways—at least for the 'classical values'. While I think that *non-truth-functional* approaches are worth exploring, e.g. [Beall, 2005a], I omit those here due to space considerations.

hence, $\neg A$. If 'Other' is to play the role for which it was introduced—namely, to correctly characterize, perhaps among other things, Liars—then one would want $\mathsf{Other}\langle A\rangle$ to be dtrue. But if, as supposed, $\mathsf{Other}\langle A\rangle$ implies $\neg\mathsf{dT}\langle A\rangle$ and, hence, implies $\neg A$, then inconsistency abounds: from the fact that A is *other*, it follows that A is *not dtrue*, in which case, since that is precisely what A says, A is also dtrue, by the essential intersubstitutivity of dtruth.

In general, then, it is not difficult to see that 'Other' in EC cannot be consistently understood as implying *not dtrue*.[51] The point applies in particular to the Strong Kleene proposal of Kripke: it makes no sense to say of Liars that they are neither dtrue nor dfalse.[52]

In the end, while the account provides a (paracomplete) response to NTP, Kripke's Strong Kleene proposal affords no answer to ECP.[53] As a result, there's no 'significantly semantic predicate' that is introduced—in the given language (as opposed to model-relative metalanguage-terms)—for purposes of correctly characterizing Liars, at least on Kripke's paracomplete account. And this is thought by some, perhaps many, to be an inadequacy of the proposal.

I will return to the issue of ECP below. For now, I turn to one more issue concerning Kripke's Strong Kleene proposal, the issue of a suitable conditional.

Dtruth and dT-biconditionals

What are standardly called T-biconditionals, biconditionals of the form $\mathsf{T}\langle A\rangle \leftrightarrow A$, have long been thought to be an essential feature of truth, something at least essential to 'the' so-called naïve theory of truth. But what about *dtruth*? What is essential to dtruth is its transparency, its intersubstitutivity. Whether all instances of $\mathsf{dT}\langle A\rangle \leftrightarrow A$ hold depends entirely on the sort of conditional one has, whether, for example, one's conditional $\rightarrow$ is such that $A \rightarrow A$ is valid. If $A \rightarrow A$ is valid in the language, then (obviously) the dT-biconditionals will thereby hold (assuming, as I will, that $\leftrightarrow$ is defined as usual).

Whether an acceptable theory of dtruth must be one that validates all dT-biconditionals is an important question that I'll none the less leave open. For present purposes, I will assume—without argument—that, other things being

[51] Recall that we are focusing on *dtruth*, which requires the intersubstitutivity of $\mathsf{dT}\langle A\rangle$ and A. Deviating from dtruth affords more options, and in many ways is what (or, more accurately, what may have) motivated approaches like [Feferman, 1984], [Maudlin, 2004], and so on.

[52] Recall the difference between the *model language* and the *real language*, and in particular the difference between *model-relative* (ultimately, merely instrumental) notions and the 'real notions' intended to be modeled. It makes perfect sense, *standing squarely and only in the classical fragment (metalanguage)*, to say of sentences in the model-language that they're neither true-in-the-model language nor false-in-the-model language. But such model-relative notions aren't at issue here.

[53] In fact, one might take some of Kripke's remarks to suggest that, by his lights, there's no escaping a 'Tarskian' or 'hierarchical' approach to ECP. (See the famous passage about Tarski's ghost [Kripke, 1975].) But I will not dwell on the exegetical issue here. [Note that one might take the 'classical reading' of Kripke, formulated by Feferman's KF, as having an answer to ECP, and in many respects that's correct. (One can truly say, e.g., that Liars are not true.) But, again, KF gives up dtruth (and, so, NTP wrt dtruth).]

equal, an account of dtruth that validates all dT-biconditionals is prima facie more attractive than an account that fails to do so. Accordingly, an account of dtruth for which we have the validity of $A \rightarrow A$ is prima facie more attractive than an account for which $A \rightarrow A$ isn't valid (at least other things being equal).[54]

That $A \rightarrow A$ is not valid in Kripke's Strong Kleene proposal is clear, since $\rightarrow$, in Strong Kleene, is simply the material conditional, which is defined $\neg A \vee B$, and here abbreviated as $A \supset B$. But, then, $A \supset A$ is valid only if $\neg A \vee A$ is valid, that is, only if LEM is valid. But LEM isn't valid in Strong Kleene, as noted in §4.1. Indeed, the heart of paracomplete proposals is the rejection of LEM.

While Kripke's proposal seems to show how we can have a non-trivial (indeed, consistent) dtruth predicate despite the existence of Liar-like sentences, it fails to show how we can achieve as much in a language for which the dT-biconditionals hold.

One might think it an easy fix to add a conditional. After all, Łukasiewicz's (pronounced 'woo-kush-YE-vitch') 3-valued language differs from the Strong Kleene language only in that it adds a conditional for which $A \rightarrow A$ is valid. (One can retain the hook, of course, so as to have two conditionals, $\rightarrow$ and $\supset$. This is no surprise given that, as noted above, $A \supset B$ simply is $\neg A \vee B$.) Łukasiewicz's conditional is defined thus:[55]

$\rightarrow_3$	1	n	0
1	1	n	0
n	1	1	n
0	1	1	1

As one can see, $A \rightarrow_3 A$ is always designated in the Łukasiewcz semantics, and hence, given the essential intersubstitutivity of dtruth, $\mathsf{dT}\langle A\rangle \leftrightarrow_3 A$ is always designated (with $\leftrightarrow_3$ defined as usual via conjunction).

The *trouble*, however, is that this proposal will not work in the sort of paracomplete, fixed-point languages at issue. One way to see this is to consider a version of *Curry's paradox*.[56] Without getting into the technical details, a simple way to see the problem is via an informal Curry-like situation. Assume a Strong Kleene (fixed point) language augmented with the Łukasiewicz conditional above. Let A be a sentence that says $\mathsf{dT}\langle A\rangle \leftrightarrow_3 \bot$, where $\bot$ is some false sentence in the semantic-free fragment, say, '$1 = 0$'. Let C be a sentence that says $\mathsf{dT}\langle C\rangle \leftrightarrow_3 \mathsf{dT}\langle A\rangle$. A paracomplete theorist will want to say that A is to receive value $\frac{1}{2}$ (or modeled as such). Suppose that C receives the value $\frac{1}{2}$. Then the values of C and A are

[54] One natural route towards an *argument* is given by Feferman [1984], who argues that without the validity of $A \rightarrow A$ 'ordinary reasoning' is crippled.

[55] Here, $n = \frac{1}{2}$.

[56] See [Curry, 1942] and, for relevant discussion, [Meyer *et al.*, 1979]. Note: Some authors [Barwise and Etchemendy, 1987] call Curry's paradox *Löb's paradox*, mostly due to the similarity between Löb's Theorem (or the proof thereof) and Curry's paradox. (The situation is somewhat similar to Gödel's Incompleteness proof and the Liar paradox.) But Curry certainly discovered the paradox much earlier—and, indeed, it's likely that Medieval logicians discovered it prior to Curry.

the same, in which case C gets the value 1. Contradiction. Similarly, a contradiction arises if C receives 0 or, obviously, 1. Hence, there's no obvious way to add Łukasiewicz's conditional to the Strong Kleene (fixed point) language.[57]

Curry's paradox, in general, imposes constraints on conditionals. A *Curry sentence* is one that says of itself (only) that *if* it is dtrue, then everything is dtrue.[58] So, a Curry conditional is a sentence C equivalent to $C \rightarrow \bot$, where $\bot$ is 'explosive', a sentence implying triviality (e.g., 'Everything is dtrue'). Let us assume that a 'genuine conditional' is one that *detaches* in the following form (sometimes called 'rule-form'), where $\Vdash$ is our consequence relation (either semantically or proof-theoretically defined).

$$A, A \rightarrow B \Vdash B$$

And assume that we want a genuine conditional such that $A \rightarrow A$ is valid (and, hence, given dtruth, the dT-biconditionals are valid). Then to avoid Curry paradox, one must either reject *reductio* or *contraction* in the following (among other) forms.[59]

R. $A \rightarrow \neg A \Vdash \neg A$

C1. $A \wedge (A \rightarrow B) \rightarrow B$

C2. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$

C3. $A \rightarrow (A \rightarrow B) \Vdash A \rightarrow B$

Consider, for example, (C1). Assume that we have a conditional $\rightarrow$ such that $A \rightarrow A$ is valid (and, hence, the dT-biconditionals), and also have both dT and Curry sentences in the language. Let C be such a sentence, equivalent to $\mathsf{dT}\langle C\rangle \rightarrow \bot$, in which case, via the dT-biconditionals and intersubstitutivity, we have

$$\star \qquad C \leftrightarrow (C \rightarrow \bot)$$

[57]This is slightly misleading, as put, but without getting into a much more technical discussion, I will leave the general problem there.

[58]The consequent of a Curry conditional may be any sentence that implies triviality or near-enough triviality (e.g., '$1 = 0$'), where a language (or truth predicate) is trivial iff everything is true (in the language).

[59]I should also note that in *paracomplete* languages in which $A \rightarrow A$ is valid and $\rightarrow$ is 'genuine' (and one has a dtruth predicate), one must also reject Importation and 'Introduction' in the following forms.

Imp. $A \rightarrow (B \rightarrow C) \Vdash A \wedge B \rightarrow C$

Int. From $A \Vdash B$ infer $\Vdash A \rightarrow B$

The import of rejecting (Int), so put, is that 'deduction theorems' won't generally hold. (I briefly discuss a related issue in §6.) It's worth noting that the otherwise apparent obviousness of (Imp) rests largely on assumptions of contraction-principles that, for Curry reasons, must be rejected. A nice discussion of these issues can be found in 'relevant logic literature' [Anderson and Belnap, 1975; Anderson *et al.*, 1992; Routley *et al.*, 1982] and, of direct relevance here, Field's [2002; 2003a]. See too [Restall, 2000].

But, then, from (C1) we get

$$C \wedge (C \rightarrow \perp) \rightarrow \perp$$

which, via substitution, yields

$$C \wedge C \rightarrow \perp$$

which, assuming normal conjunction, is equivalent to

$$\star\star \quad C \rightarrow \perp$$

Given that $\rightarrow$ is a 'genuine conditional' and, so, detaches, we quickly get C itself via ($\star$), and now detaching again via ($\star\star$) yields $\perp$. Triviality.

One might think that, while his 3-valued conditional won't work, the full continuum-valued language $\mathcal{L}_\infty$ of Łukasiewicz, which, except for more values, retains the (K1)–(K3) clauses for standard connectives (see §4.1), might do the trick. In particular, $\rightarrow_\infty$, the conditional in $\mathcal{L}_\infty$, satisfies none of the given contraction principles nor (R).[60] Unfortunately, the proposal won't work, as Greg Restall [1992] and, more generally, [Hajek *et al.*, 2000] showed: the resulting theory will be ω-inconsistent.[61]

4.2 Field

The upshot of §4.1 is that while Kripke's paracomplete proposal shows that, despite having Liars in the language, we can have a non-trivial (indeed, consistent) dtruth predicate (in the language, for the language), it none the less exhibits two apparent inadequacies.

- I1. The proposal fails to answer NTP for a language in which all dT-biconditionals hold.
- I2. The proposal fails to answer ECP in any fashion.

As in §4.1, Curry's paradox puts constraints on adding a genuine conditional for which $A \rightarrow A$ is valid (and, hence, for which the dT-biconditionals all hold). Answering ECP, in turn, likewise requires care, since 'revenge Liars' are ever-ready to emerge.

Hartry Field, in a series of papers (see throughout), advances the Kripkean paracomplete proposal by attempting to overcome (I1) and (I2), the two notable inadequacies of Kripke's own proposal. Field maintains the basic paracomplete line that paradoxes (perhaps among other phenomena) teach us that LEM is to be

[60] More promising yet, perhaps, is that naïve 'property theory' (sometimes called 'naïve set theory') is consistent in $\mathcal{L}_\infty$. See [White, 1979].

[61] I will not get into the details of formulating, e.g., Peano Arithmetic in $\mathcal{L}_\infty$ and showing the resulting ω-inconsistency. For details, I recommend beginning with [Restall, 1992] and, in turn, the more generalized [Hajek *et al.*, 2000].

rejected, that some instances of $A \vee \neg A$ are to be rejected.[62] Field's contribution is an answer to NTP for a language in which all dT-biconditionals hold and, in turn, an answer to ECP.[63]

Field's main contributions, as above, are his conditional and his approach towards characterization (classifying Liars, etc.), with the latter nicely falling out of the former.[64] Given the aims of this essay, I will only provide a sketch of (the basic idea of) the conditional and, in turn, Field's approach towards characterization.[65] But before sketching the conditional, I will first (briefly) discuss the background philosophical picture, and then turn to (a sketch of) the formal model.

Philosophical picture: stronger truth

An inadequacy of Kripke's proposal is that we're left with nothing to truly say about Liar-like sentences, at least not *in our language* (the language that enjoys its own dtruth predicate).[66] Intuitively, the paracomplete theorist thinks that, for purposes of truly characterizing or 'classifying' Liar-like sentences, we need to acknowledge an additional 'semantically significant' category beyond dtruth (and dfalsity). But how shall this 'other' category be understood?

As in §4.1, there's no clear sense in saying that Liars are *neither dtrue nor dfalse.* Still, one might think that there's *some* sense in which Liars are 'not true' or 'not false'. This thought motivates Field's proposal.[67]

---

[62]This is not to say that there aren't significant (proper) fragments of the language for which LEM holds. For example, one might maintain that LEM holds over the arithmetical fragment of our language, or physics, or so on. (Field [2003b] explicitly agrees with this, suggesting that LEM may hold for mathematics or the like, in general.) Of course, as mentioned in §4, since we're (here) dealing with dtruth, *rejection of* A is not going to be *acceptance of* $\neg A$. Field naturally suggests weakening the classical 'exhaustion' rule for probability (or degrees of belief), namely, $Pr(A) + Pr(\neg A) = 1$, to one that respects a paracomplete approach: namely, $Pr(A) + Pr(\neg A) \leq 1$. There are subtle issues here, but for space I leave them aside.

[63]The respect in which Field answers the intended import of (the admittedly imprecise) ECP is something I'll briefly discuss below. What I should emphasise is that, in my opinion, Field's work is a remarkable advance in the area of *dtruth* and paradox. (And, of course, Field has also contributed a great deal to the philosophical issues concerning *dtruth.* See [Field, 2001].)

[64]Actually, Field should also be given credit for explicitly noting that paracomplete accounts of rejection (and acceptance) call for an adjustment of classical probability theory. While this point is often assumed, Field took the time to explicitly propose a revision, one that, by my lights, seems to be entirely right (for purposes of a paracomplete account). See [Field, 2003b] for a nice discussion.

[65]Field continues to revise the proposal, coming up with further refinements and improvements. My aim is not to give the very latest, or even the full details of any particular version. Rather, my aim is merely to sketch the basic idea.

[66]Kripke precisely defines notions of *ungroundedness* and *paradoxical* (and more). One might think that we can truly characterize Liar-like sentences as *paradoxical*, so defined (per Kripke). But so defined, one can't consistently add 'is paradoxical' to the language, at least if it is to behave as Kripke seems to suggest—e.g., as a predicate the extension and anti-extension of which are exhaustive and exclusive. It is precisely that assumption that paracomplete theorists ought to reject! And Field recognizes as much.

[67]The thought likewise seems to motivate a related (and well-known) earlier proposal by Vann McGee [1991]. McGee's work precedes that of Field's, but the two are related.

The paracomplete theorist rejects that Liars are dtrue or not dtrue (or, equivalently, dfalse or not). But perhaps one can recognize a stronger notion of 'truth' according to which Liars and their negations are *not true*. Let sT be our 'stronger truth predicate' (stronger than dT). Being a stronger notion than dtruth, one might have it that while both

st1. $\Vdash \mathsf{sT}\langle A\rangle \rightarrow A$ for all A (and for some suitable conditional!)

st2. $\mathsf{sT}\langle A\rangle \Vdash A$ for all A

hold, the converse of either (st1) or (st2) fails.[68] This failure need not get in the way of expressing generalizations or the like, since *dtruth* remains as before—full intersubstitutivity holds. Dtruth remains our expressive device. The new device, sT, is brought in to do a job that dtruth was never intended to do: namely, *fail to be transparent*! In particular, we want a device that allows us to truly (dtruly) 'classify' Liars, sentences like the ticked sentence in §3. For reasons above, we can't classify the ticked sentence or its negation as dtrue (or, hence, dfalse); however, with our stronger notion, we may be able to classify the sentence and its negation as *not strongly true*. That is the idea.[69]

Care, of course, must be taken. Obviously, on pain of 'revenge paradox', sT, however it is spelled out, must be such as to resist excluded middle, resist having $\mathsf{sT}\langle A\rangle \vee \neg\mathsf{sT}\langle A\rangle$. (Just consider a sentence A equivalent to $\neg\mathsf{sT}\langle A\rangle$.) And this is precisely in keeping with the paracomplete theorist's rejection of LEM: that there is *no* 'truth-like' or 'negation-like' device † (be it a predicate or operator) in the language that is 'exclusive' in the sense of satisfying both of the following.

e1. $\Vdash \dagger A \vee A$ for all A.

e2. $\dagger A, A \Vdash B$ for all A and B.

The heart of paracomplete accounts is that at least (e1) fails for our language.[70]

But now an apparent tension arises. The reason that we want a stronger notion of truth is that we want to be able to dtruly classify Liar-like sentences, 'characterize' them as being *in some sense* 'not true'. So, we bring in sT, and we want to say for any Liar-like sentence A in the sT-free fragment that A is not strongly true; we want to assert $\neg\mathsf{sT}\langle A\rangle$. This much is not difficult, provided we're restricting ourselves to As in the sT-free fragment. The *trouble*, of course, is that we want to talk about *any* sentence in the *full* language, including any sT-ful sentences—sentences that use 'sT'. But, then, we want to be able to say of sT-ful Liars that they are not strongly true. And that's the problem. Inevitably,[71] there will be

[68] Typically, e.g. [McGee, 1991], the converse of (st1) fails. This is likewise the case for Field's proposal.

[69] Such 'strong truth' is often understood in terms of *determinately dtrue*, the idea being that by rejecting each of A and $\neg A$, one is rejecting that either sentence is determinate.

[70] Note that we're not talking about *model-relative* notions that may be defined in an entirely classical, *proper* fragment of one's full, paracomplete language. We're talking about 'non-model-relative' or, as Field (in conversation) says, 'absolute' notions.

[71] Formally, the force of 'inevitably' is ensured by Gödel's diagonal lemma.

sT-ful sentences A such that, were either $\mathsf{sT}\langle A\rangle$ or $\neg\mathsf{sT}\langle A\rangle$ to hold, inconsistency would follow. Avoiding such inconsistency requires, as above, the rejection of (the equivalent of) excluded middle for sT. But, now, there would seem to be sT-ful sentences that fail to be 'correctly classified' by sT.

The apparent tension can be seen as follows.[72] We *begin* with a transparent expressive device dT, which, when introduced into the language (to play its transparency role), gives rise to dT-ful Liar-like sentences. The paracomplete theorist maintains that such sentences are not problematic if we reject LEM, and in particular (at least) the Liar-instances of LEM. So, we can keep our (consistent) expressive device dT despite its inevitable Liars. One problem solved. But, *next*, we want to be able to dtruly 'classify' or 'characterize' the given Liars. Towards that end, we introduce a stronger notion of truth, sT.[73] And now we can use sT to dtruly classify all dT-ful Liars. Another problem solved. But *now again* to avoid inconsistency, we must likewise reject LEM for sT, and in particular reject sT-ful Liar-instances of LEM. We now seem to require *yet another* 'even stronger truth' to dtruly classify the given sT-ful Liar-sentences (ones remaining 'unclassified' on pain of inconsistency). And so on, ad infinitum.

And now the tension is clear. Suppose that we have some *unified predicate* (say, 'Other') that characterizes all Liar-like sentences, so that we have the following.

EC. $\Vdash \mathsf{dT}\langle A\rangle \vee \mathsf{dF}\langle A\rangle \vee \mathsf{Other}\langle A\rangle$

Then we seem to be stuck in inconsistency, at least assuming normal behavior for the (extensional) connectives.[74] After all, consider a sentence A that says $\neg A \vee \mathsf{Other}\langle A\rangle$. By EC, A is either dtrue, its negation dtrue, or it is Other. If dtrue, then A is dfalse or Other. If dfalse or Other, then A is dtrue. If, then, the given semantically significant 'categories' are exhaustive and exclusive, then inconsistency arises.[75]

What the picture of infinitely many 'stronger truth predicates' requires is the absence of any unified (semantically significant) predicate in terms of which all Liar-like sentences are to be classified.[76] The given picture must be accompanied by a rejection of anything yielding something along the lines of EC.

Perhaps the rejection of EC—the rejection of a unified (semantically significant) predicate in terms of which all Liar-like sentences are classified—is not unnatural. After all, the heart of standard paracomplete accounts is a rejection of LEM *all the*

[72] I will assume a Strong Kleene approach to the basic, extensional connectives. This is the case with all paracomplete accounts that I discuss—Kripke, Field, and the 'paranormal' suggestion.

[73] Here, sT can be either a predicate or operator. The difference is minimal when our basic truth predicate is dtruth.

[74] But see §4.3.

[75] Again, unfortunately I remain vague about what counts as 'semantically significant', and I believe that the project of clarifying this notion is an important (and pressing) one. For now, one can think of 'dtrue' and 'dfalse' among the lot, and then any predicates in terms of which Liars are to be 'classified'.

[76] Note that this immediately requires rejecting the coherence of 'quantifying over the hierarchy' of such predicates in the—otherwise intuitive—sense of, e.g., *true in some sense or other of 'strongly true'* or the like.

way through. In broadest (though, admittedly, somewhat vague) terms, the rejection of LEM might be seen as a basic rejection of an 'exhaustive characterization' in terms of unified (semantically significant) predicates.

The basic tension, in the end, is one arising from a common aim: the aim of dtruly classifying (within one's language) all Liar-like sentences in one's given language.[77] Intuitively, the aim is (presumably) to achieve such exhaustive characterization using 'unified' semantically significant predicates, something that yields the likes of EC. But another route is available: classifying any given Liar via infinitely many 'stronger and stronger' truth predicates, none of which afford a 'unified predicate' that, as it were, serves as a 'unionizer' of all such predicates. While Field does not achieve the former, he none the less provides a powerful approach along the latter lines.

Formal model: conditional and determinacy

As above, Field aims to retain a consistent dtruth predicate but, going beyond Kripke, also have all dT-biconditionals *and* a way of 'classifying' any given Liar-like sentence in the language. Field shows how to add a suitable conditional to the Kripkean Strong Kleene framework, and then defines 'stronger truth' or, as Field says, 'determinate truth', in terms of the given conditional. Given that, as in §4.2, any such 'stronger truth' predicate (or operator) must resist LEM, Field's aim of 'characterizing' Liars requires infinitely many such 'stronger truth' devices. One notable feature of Field's framework is that the requisite infinite stock of (stronger and stronger) truth devices falls out of Field's constructed conditional.

This section presents only a sketch of Field's basic idea for introducing a suitable conditional into a paracomplete—and otherwise merely Strong Kleene—language with a (consistent) dtruth predicate. I first present an initial sketch of (the basic idea of) how to extend Kripke's initial construction with a (non-truth-functional) suitable conditional—what Field calls a *restricted semantics.*[78] In turn, I sketch a more general setting (what Field calls 'General semantics') for the conditional.[79]

Restricted semantics

Let K_3 be the Strong Kleene logic. Field's aim is to give an extension of K_3 that, in addition to containing a consistent dtruth predicate, validates all dT-biconditionals in such a way that Curry paradox is avoided. The basic proposal is a novel

[77]Of course, we're talking about languages that have their own, non-trivial (indeed, in this section, consistent) dtruth predicate.

[78]Strictly speaking, Field [2003a] does show how one can expand the values of the language in such a way as to view the given conditional as 'truth-functional', but I will ignore that here.

[79]I should note (again) that Field has modified his account in various ways, with the latest [2005c] being a variation of Stephen Yablo's work [2003]. I would've liked to discuss Yablo's work, in general, as well as Field's work in more detail; however, space limitations prohibit doing so. (Field's work is cited throughout. For Yablo's work, see especially [Yablo, 1993a; Yablo, 1993b].)

combination of ideas from Kripke [1975] and *revision theorists* [Herzberger, 1982; Gupta and Belnap, 1993].

We start with a (first-order) syntax supplemented with dT and a primitive two-place connective $\rightarrow$ (which I'll call *the conditional*). (Any sentence the main connective of which is the conditional will be called a *conditional.*) With respect to the conditional-free fragment, the language is interpreted exactly along the lines of Kripke (as in §4.1). The challenge is to interpret all sentences, including all conditionals, in such a way as to retain a (consistent) dtruth predicate—and, so, achieve full intersubstitutivity even with respect to (and 'inside of') conditionals—and validate all dT-biconditionals.

Field's proposal is to interpret the language via a transfinite sequence of Kripkean (Strong Kleene) fixed points $\mathcal{P}^{\alpha}$ (for 'point α', with α an ordinal),[80] where each such fixed point is 'built from' an initial *starting valuation* $\mathcal{S}_{\alpha}$ (for 'start α'), which assigns elements of $\{1, \frac{1}{2}, 0\}$ to *all and only condtionals.* Beginning with such 'start points', Kripke's construction (see §4.1) yields a value for every sentence in the language in such a way that dtruth (transparency, intersubstitutivity) is preserved. With respect to such start points $\mathcal{S}_{\alpha}$, and in particular how any given $\mathcal{S}_{\alpha}$ is determined on the basis of 'prior' Kripkean $\mathcal{P}^{\beta}$, Field proposes the following recipe.[81]

Fb. Base (Zero). $\mathcal{S}_0(A \rightarrow B) = \frac{1}{2}$ for all A and B.

Fs. Successor. At successor points (or stages), we look back at the prior Kripkean fixed point:

$$\mathcal{S}_{\alpha+1}(A \rightarrow B) = \begin{cases} 1 & \text{if } \mathcal{P}^{\alpha}(A) \leq \mathcal{P}^{\alpha}(B) \\ 0 & \text{otherwise.} \end{cases}$$

Fl. Limit. At limit points (stages), we look backwards at all prior Kripkean fixed points:

$$\mathcal{S}_{\lambda}(A \rightarrow B) = \begin{cases} 1 & \text{if } \mathcal{P}^{\gamma}(A) \leq \mathcal{P}^{\gamma}(B) \text{ for some } \beta < \lambda \\ & \quad \text{and any } \gamma \text{ such that } \beta < \gamma < \lambda \\ 0 & \text{if } \mathcal{P}^{\gamma}(A) > \mathcal{P}^{\gamma}(B) \text{ for some } \beta < \lambda \\ & \quad \text{and any } \gamma \text{ such that } \beta < \gamma < \lambda \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

So goes the construction of 'start points' from the 'prior' Kripkean fixed points. As above, the latter—the Kripkean fixed points—are the points that yield the 'ultimate values' in terms of which all sentences eventually stabilize into a language with both dtruth and all dT-biconditionals. The various $\mathcal{P}^{\alpha}$, as above, are determined entirely by the $\mathcal{S}_{\alpha}$ (which give values to the conditionals) and the Strong

[80]Note that these are fixed points of Kripke's 'jump operator', which will henceforth be left as implicit.

[81]For ease, I will write '$\mathcal{S}_{\alpha}(A)$' to abbreviate *the value of A in the start point $\mathcal{S}_{\alpha}$*.

Kleene (minimal) fixed point construction—the various clauses for compounds, (K1)–(K3). (See §4.1.)

Brief reflection on (Fb)–(Fl) indicates that values assigned to (at least typically paradoxical) sentences at the various $\mathcal{S}_\alpha$ and, in turn, $\mathcal{P}^\alpha$, fluctuate quite a bit; such sentences exhibit jumpy instability. By way of settling on 'ultimate values', by way of bringing about order to such apparent chaos, Field takes a leaf from revision theory (see §8). In particular, Field defines the *ultimate value of* A, say $|A|$, thus:

$$|A| = \begin{cases} \texttt{lim}_\beta \mathcal{P}^\beta(A) & \text{if the limit exists} \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

In other words, if for some point α such that for any $\beta \geq \alpha$ we have it that $\mathcal{P}^\beta(A)$ is 1, then $|A| = 1$. Similarly for 0. Let $n \in \{1, 0\}$. The idea is simply that if A is eventually forevermore assigned n, then n is A's ultimate value. But, of course, there may be no such point beyond which the value of A 'stabilizes' at either 1 or 0, in which case $|A| = \frac{1}{2}$. More precisely, $|A| = \frac{1}{2}$ if either A is never eventually forevermore assigned anything or there's some point α such that for any $\beta \geq \alpha$ we have it that $\mathcal{P}^\beta(A) = \frac{1}{2}$ (i.e., A is eventually forevermore assigned $\frac{1}{2}$).

Field [2003a] proves that such 'ultimate values' obey the K_3-rules for extensional connectives, that is, for all connectives except the non-extensional $\rightarrow$, the conditional.[82] Moreover, he shows that the construction validates all dT-biconditionals.[83]

I will discuss the virtues of Field's proposed conditional, and its role with respect to 'stronger truth', in §4.2. For now, I turn to a more general, and perhaps philosophically more 'intuitive', account of the conditional.

General semantics

Because other approaches to dtruth (e.g., paraconsistent) often invoke a 'possible worlds' framework for purposes of modeling a suitable conditional, I will briefly sketch Field's 'general semantics' for his conditional, a semantics that is related to possible worlds approaches, though the intended philosophical interpretation, as Field [2003b] remarks, is better thought of as 'possible assignments relative to actual conditions or constraints'. In this (sub-) section I will simply sketch the construction, leaving comments to §4.2.[84]

The aim, once again, is to give an extension of K_3, but in this case we work with a 'modalized' K_3 (first-order) language, where, as above (throughout), we can stipulate that the 'semantic-free' fragment is entirely classical, and the dtruth

[82] More technically, Field's 'Fundamental Theorem' shows that there are ordinals γ such that for any (non-zero) β the value of any sentence A at $\gamma \cdot \beta$ is just $|A|$.

[83] Field also shows that the dT-biconditionals are validated in a stronger sense that he dubs 'conservative', the idea being (roughly) that the resulting theory is consistent *with any arithmetically standard starting model*. For discussion, see [Field, 2003a].

[84] A few caveats: first, the 'general semantics' was motivated largely towards a unified solution to both semantical and soritical (vagueness-) paradoxes. I will not discuss the latter here. Second, I will not be giving all of the constraints that Field proposes for purposes of achieving various desirable features (of the conditional); I will simply sketch the basic idea.

predicate is achieved along Kripkean lines, as above.[85] The difference, now, is that we expand our interpretations with a (non-empty) set $\mathcal{W}$ of 'worlds' and, in turn, assign values to sentences relative to such worlds, where the values remain either 1, 0, or $\frac{1}{2}$. This much is standard. The task is to tweak the interpretations in such a way as to give the target conditional its desired features.

Field's proposal is a novel variation on so-called 'neighborhood semantics'.[86] We let $\mathcal{W}$ be an infinite set of worlds at which sentences are assigned an element of $\{1, \frac{1}{2}, 0\}$, letting @ be a (unique) distinguished element of $\mathcal{W}$, the 'actual world'. In turn, we impose a 'similarity relation' on $\mathcal{W}$ in such a way that each $w \in \mathcal{W}$ comes equipped with a set of 'sufficiently similar worlds' (a so-called 'neighborhood' of w), worlds that satisfy some condition of similarity with respect to w. Specifically, Field proposes that each $w \in \mathcal{W}$ be assigned a (possibly empty) *directed* family $\mathcal{F}_w$ that comprises non-empty elements of $\wp(\mathcal{W})$, non-empty subsets of $\mathcal{W}$. The *directedness* of $\mathcal{F}_w$, which amounts to

$$(\forall w \in \mathcal{W})(\forall \mathcal{X}, \mathcal{Y} \in \mathcal{F}_w)(\exists \mathcal{Z} \in \mathcal{F}_w)\, \mathcal{Z} \subseteq \mathcal{X} \cap \mathcal{Y}$$

allows for 'incomparability', that is, that the relation of similarity needn't be linear.[87]

With an eye towards semantical paradox, a few other tweaks are required.[88] Define, for any $w \in \mathcal{W}$:

Normality. w is *normal* iff $w \in \mathcal{X}$ for all $\mathcal{X} \in \mathcal{F}_w$.

Abnormality. w is *abnormal* iff it is not normal.

Loneliness. w is *lonely* iff $\{w\} \in \mathcal{F}_w$.

Happiness. w is *happy* iff it is not lonely.

Field stipulates that @ be both normal and happy on any interpretation, but otherwise worlds may be abnormal and lonely.[89] Accordingly, every interpretation is such that, per normality, $@ \in \mathcal{X}$ for all $\mathcal{X} \in \mathcal{F}_@$ and, per happiness, $\{@\} \notin \mathcal{F}_@$. Hence, @ is 'sufficiently similar' to itself on all interpretations, and @ is also 'sufficiently similar' to some $w \neq @$ on all interpretations.

As above, sentences are now assigned a value at each world. With respect to the conditional-free fragment, the valuations simply follow the Strong Kleene rules;

[85] As noted above, the framework is motivated not only by semantical paradox but also soritical. As a result, if we were concentrating on vagueness, we wouldn't stipulate an entirely classical 'semantic-free' language.

[86] The generalization of standard Kripke models is also sometimes called *Montague semantics*, or *Montague-Scott* semantics. For general references, see [Chellas, 1980].

[87] Field [2003b] notes that one could, without radical deviation from the basic proposal, simply impose a linear ordering via $\subseteq$.

[88] Actually, achieving all of Field's desired features requires other constraints on interpretations, but for present purposes I skip over them.

[89] That @ is to be happy is not just warmheartedness on Field's part. If @ were allowed to be lonely, various (contraction-related) validities would emerge that would engender inconsistency in the resulting dtruth theory.

(K1)–(K3), in effect, are modified only with respect to being relativized to worlds, even though reference to worlds in the clauses for extensional connectives makes no essential difference. (In other words, where A is conditional-free, the value of A at w depends only on the values of A's constituent parts at w. One needn't 'look at other worlds' to figure out the value of purely extensional sentences.) The worlds come into play with conditionals.[90]

$$|A \rightarrow B|_w = \begin{cases} 1 & \text{if } |A|_{w'} \leq |B|_{w'} \text{ for some } \mathcal{X} \in \mathcal{F}_w \text{ and any } w' \in \mathcal{X} \\ 0 & \text{if } |A|_{w'} > |B|_{w'} \text{ for some } \mathcal{X} \in \mathcal{F}_w \text{ and any } w' \in \mathcal{X} \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

With valuation-conditions in hand, the (semantic) consequence relation $\Vdash$ may be defined. Towards that end, let us say that, relative to an interpretation, a sentence A is *actually verified* iff $|A|_@ = 1$ (in the given interpretation). Similarly, a set Σ of sentences is *actually verified* iff B is actually verified, for each $B \in \Sigma$. Then the (semantic) *validity* relation $\Vdash$ is defined thus:

$\Sigma \Vdash A$ iff any interpretation that actually verifies Σ actually verifies A

with *valid sentences* being consequences of $\emptyset$. Given the existence of abnormal and lonely worlds (or the existence of interpretations containing as much), other notions of validity may be introduced, but for present purposes I will focus just on the given notion.[91]

So goes the basic model. While, for space reasons, I have left out various (some not insignificant) details, there is enough in the foregoing to turn to philosophical discussion.[92]

Comments: virtues and strong(-er) truth

I will concentrate on the general semantics of Field's conditional, focusing on the philosophical features. Right off, one of the chief desiderata is plain.

$\Vdash A \rightarrow A$

[90] I will continue to use the bar-notation, e.g. $|A|$, to abbreviate 'the value of A'. (This follows Field's notation, and ties in the earlier discussion of 'ultimate values'.) The difference, of course, is now that values are relative to worlds, and so, e.g., $|A|_w$ is the value of A at w.

[91] In standard 'non-normal worlds' semantics, wherein one has different 'types' of worlds, broader notions are standardly introduced by various restrictions on the 'types' of worlds invoked in one's definition. Field's neighborhood account is similar, and [Field, 2003b] introduces 'universal validity' (quantifying over all worlds of all interpretations) and 'strongly valid' (all normal worlds of all interpretations). For space reasons, I skip over these distinctions.

[92] What may not be plain is how Field's 'general semantics' is a more general version of the noted 'restricted semantics'. The short answer is that the latter can be seen as a special case of the former: one allows for 'normal ordinals' (analogous to 'normal worlds') in the latter, and modifies the account of validity in the latter in terms of a distinguished such 'normal ordinal' (something guaranteed by Field's Fundamental Theorem). For discussion see [Field, 2003b].

Hence, given the intersubstitutivity of dT, which is preserved in *all* (transparent) contexts, all dT-biconditionals are similarly valid. And since Field [2003a] gives a consistency proof for the resulting language (and dtruth theory), this amounts to a remarkable step forward in (consistent) paracomplete accounts of dtruth. Kripke provided an answer—a paracomplete answer—to NTP for a language in which not all dT-biconditionals hold. What Field's conditional has given us is an answer—a paracomplete answer—to NTP for a language in which all dT-biconditionals hold, an answer that preserves the insights of Kripke but goes further, properly extending the resulting logic.

But there's more. The given conditional also exhibits various familiar features.[93]

$A, A \rightarrow B \Vdash B$ $\qquad$ $A, \neg B \Vdash \neg(A \rightarrow B)$

$\Vdash \neg\neg A \rightarrow A$ $\qquad$ $\Vdash A \wedge B \rightarrow A$

$\Vdash \neg(A \rightarrow B) \rightarrow (A \vee \neg B)$ $\qquad$ $\Vdash (A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$

$\Vdash \forall x A \rightarrow A(t/x)$ (for proper substitution)

In addition to other features that are here omitted (for space reasons), Field's conditional exhibits fairly natural behavior, at least within a general paracomplete setting in which LEM is rejected (which, of course, is the aim). In particular, $\rightarrow$ behaves like $\supset$ when excluded middle is assumed, and so we have, for example,

$$A \supset B \Vdash A \rightarrow B$$

and, indeed,

$$A \vee \neg A, B \vee \neg B \Vdash (A \supset B) \leftrightarrow (A \rightarrow B)$$

and similarly for material equivalence.[94]

In the remainder of this (sub-) section I will briefly address a few of the pressing philosophical issues mentioned above: Curry, strong (-er and stronger) truth, and, relatedly, 'exhaustive characterization'.

Curry paradox

Of course, given Field's consistency proof [2003a], one knows that Curry paradox doesn't pose problems; however, it is worth (at least very briefly) touching on the issue—and so I will, but only very briefly.

I've assumed, as above, that a minimum requirement on a 'genuine conditional' is that it detach in the sense that the argument from $\{A, A \rightarrow B\}$ to B is valid. Field's conditional is genuine. But Field's conditional also validates Identity, $\Vdash A \rightarrow A$, so that (given dtruth) the dT-biconditionals all hold. In such

[93] For a list of other notable features see [Field, 2003a; Field, 2003b; Field, 2005c], and for a very illuminating discussion see [Yablo, 2003].

[94] Note well: These last features depend on having other features that I've here omitted, features that require other constraints being imposed on the 'neighborhood interpretations'. With suitable constraints imposed, e.g., one gets the 'meta-rule' that if arguments A/B and B/C are valid, then so too is the argument $A \vee B/C$. (This does not hold in K_3, but does hold in a proper extension of it, sometimes called K_3^+.) See Field's cited work for a full discussion.

a setting (given normal behavior of other connectives), Curry demands—on pain of triviality—giving up Reductio in form (R), contraction principles (C1)–(C4) and the like, as well as (Imp) and (Int). (See §4.1, page 206ff.) That Field's conditional does as much is not only a virtue but, as said, a necessity—on pain of triviality. Since Curry's paradox turns on the validity of such principles or rules, any argument towards Curry paradox is blocked as invalid.[95]

It might be useful to see a 'counterexample' to some of the invalid principles. Consider principle (C1), which is sometimes called *Assertion* or *Pseudo Modus Ponens*. Notice that if @ could be lonely, there would be no way to invalidate (C1). But Field's demand that @ be happy provides an immediate counterexample: just consider an interpretation in which $\mathcal{F}_@ = \{\{@, w\}\}$ and, for simplicity, let $\mathcal{F}_w = \emptyset$. Let $|A|_@ = |B|_w = |A|_w = \frac{1}{2}$ and $|B|_@ = 0$. Then $|A \rightarrow B|_@ = \frac{1}{2}$ and, hence, $|A \wedge (A \rightarrow B)|_@ = \frac{1}{2}$, and so $|A \wedge (A \rightarrow B)|_@ > |B|_@$. But most importantly, since not every world in the (unique) @-neighborhood is such that $|A \wedge (A \rightarrow B) \rightarrow B| \leq |B|$ or $|A \wedge (A \rightarrow B) \rightarrow B| > |B|$, we have what paracomplete theorists will naturally want with respect to Curry-instances of (C1).

$$|A \wedge (A \rightarrow B) \rightarrow B|_@ = \frac{1}{2}$$

And since $\frac{1}{2}$ is not designated, (C1) is invalid. Similar (and, indeed, even simpler) counterexamples are available for the other principles and rules.

Strong truth: determinacy

Field's aim is not only to validate all dT-biconditionals but also go beyond the Kripkean framework with respect to characterizing Liar-like sentences. Towards that end, Field proposes to recognize a stronger notion of truth than mere (entirely transparent) dtruth.

As mentioned in §4.2, there is always a risk of introducing more 'truth-like' devices (predicates, operators): paradoxical sentences are always ready to spring up. This is where Field's consistency proof [2003a] for the full conditional-ful (and dT-ful) language comes into play.

The consistency proof shows that the conditional doesn't introduce any further paradoxes that aren't already resolved by the guiding, paracomplete rejection of LEM. And that is the key. Field wants to characterize all Liar-sentences via 'stronger and stronger truth', and do so without bringing about yet further paradox. Given the consistency proof, it is natural to seek an account of such 'strong truth' that *invokes only the resources of the language at hand*, the language for which we have Field's consistency proof. A remarkable feature of Field's framework

[95] I should note that there are various versions of Curry's paradox, not each of which assumes *all* of the 'bad' principles or rules, but each assumes one or more of them. I should also note that, at least in paracomplete settings, one need not give up the *structural* contraction rules, as one does in paraconsistent settings. (Giving up the structural rules prohibits, in effect, 'using a premise more than once' in a proof. Field's framework preserves structural contraction, so understood.)

is that he enjoys just such an account: he defines infinitely many 'determinately' operators—these amount to 'stronger and stronger truth'—*out of the conditional.*

There are two items that need to be explained (or, at least, sketched): what is the definition of 'determinately' in terms of the conditional? How do we get infinitely many? I will (very) briefly sketch the answer to each question, leaving details to Field's work (cited throughout).

The basic account. Where $\top$ is any logical truth, Field proposes to 'introduce' a *determinately* operator $\mathbb{D}$ thus:

$$\mathbb{D}A =_{df} (\top \rightarrow A) \wedge A$$

This immediately gives standard behavior for 'determinately' operators, in particular,

d1. $A \Vdash \mathbb{D}A$

d2. $\mathbb{D}A \Vdash A$

d3. $\Vdash \mathbb{D}A \rightarrow A$

As mentioned in §4.2, standard approaches to 'determinate truth' usually give up the converse of (d3), and the situation is no different with Field. That the converse does fail may be seen by considering an interpretation in which $|A|_@ = \frac{1}{2}$, in which case $|\top \rightarrow A|_@$ is not designated, regardless of $\mathcal{F}_@$.

With $\mathbb{D}$ at hand, one can now classify $\mathbb{D}$-free Liars as proposed: neither they nor their negations are *determinately* true. For any such sentence A, we may assert $\neg\mathbb{D}\mathsf{dT}\langle A\rangle \wedge \neg\mathbb{D}\neg\mathsf{dT}\langle A\rangle$ or, equivalently given dtruth, $\neg\mathbb{D}A \wedge \neg\mathbb{D}\neg A$.[96]

But what of weaker Liars? Now that we have $\mathbb{D}$, so understood, in the language, we inevitably get $\mathbb{D}$-ful Liars—for example, sentences that say of themselves (only) that they're not determinately true, or not determinately determinately true, or so on. Such sentences call for 'stronger and stronger truth'. But Field's construction already yields as much.

Infinitely many. The point is fairly obvious. Consider a Liar L that says $\neg\mathbb{D}L$. That Field's construction handles L, so understood, follows from his consistency proof: such a sentence receives an interpretation in the language, namely, $\frac{1}{2}$. Of course, given (d1)–(d3), one cannot truly classify L as being not determinately dtrue. But one can generalize Field's proposal in a natural way: one can truly classify L as *not determinately determinately dtrue*, that is, one may assert $\neg\mathbb{D}\mathbb{D}L$.

In general, Field's operator may be iterated into the transfinite: for some suitable ordinal notation that yields σ, we have an operator $\mathbb{D}^\sigma$, the σ-many iteration of $\mathbb{D}$. And for each such 'determinately' operator, there will be (increasingly weaker)

[96] Recall that Field's construction preserves dtruth, that is, preserves the essential intersubstitutivity of $\mathsf{dT}\langle A\rangle$ and A in *all* (transparent) contexts, including 'inside' conditionals—and, hence, 'inside' $\mathbb{D}$-contexts.

$\mathbb{D}^{\sigma}$-ful Liars, each of which gets 'correctly characterized' by a stronger operator $\mathbb{D}^{\sigma+1}$ (provided the ordinal notation yields as much).[97]

What is central to the proposal is its thoroughgoing paracompleteness. LEM does not hold even for 'determinate truth'. Consistency is purchased by such thoroughgoing paracompleteness: for any $\mathbb{D}^{\sigma}$-ful Liar L_{σ} (for suitable σ), the failure of $\mathbb{D}^{\sigma}L_{\sigma} \vee \neg\mathbb{D}^{\sigma}L_{\sigma}$ will arise. Of course, as mentioned in §4.2, the failure of LEM for one's level-σ 'strong truth device' requires having an even stronger device if one wants to classify level-σ Liars. But the point is that Field's basic construction provides as much. For any Liar constructible in the language (or, at least, the hierarchy of determinately operators), there's a 'strong' or 'determinate' device that classifies the sentence.

And to repeat: there is no threat of 'determinately'-Liars wreaking havoc, since all such Liars are constructible only in a language (the full conditional-ful language) that enjoys a consistency proof. Needless to say, this is a significant improvement on the Kripkean picture.

Exhaustive characterization?

Here, the issue is delicate. What Field provides is a way of consistently classifying all Liar-like sentences (would-be paradoxical sentences) definable in the 'hierarchy of determinately operators'. This is a major step forward on other paracomplete accounts. On the other hand, one might wonder whether Field has achieved 'exhaustive characterization' in the target (but, alas, as yet imprecise) sense. Can we, for example, dtruly assert that every sentence is dtrue, dfalse, or not determinately dtrue in any (of the infinite) sense(s) of 'determinately'? Better (but still not ideally) put, is there some unified 'semantically significant' predicate Other such that we can dtruly say the following?

$$\mathsf{dT}\langle A\rangle \vee \mathsf{dF}\langle A\rangle \vee \mathsf{Other}\langle A\rangle$$

The answer is No, at least if Other carries the intuitive sense of *not determinately dtrue or determinately not dtrue, in some sense of 'determinately' or other.* In that respect, Field's proposal, while a significant advance on paracomplete approaches, seems to fail to achieve the target sense of 'exhaustive characterization' involved in ECP, imprecise as that target sense remains.

But the issue, as mentioned, is delicate. Such a 'characterization' presupposes the existence of some 'absolute determinacy operator' that, in effect, would be (like) the union of all determinacy operators. Given Field's consistency proof, such an operator simply doesn't exist, at least if our 'real language' is relevantly like Field's model language. One might, of course, say that Field's model is inadequate precisely on that score: our 'real language', unlike Field's model language, has the resources to express just such a unified 'determinacy' operator, one that is entirely intelligible. But Field [2005b] questions such alleged intelligibility.

[97]For limit ordinals, one can mimic infinite conjunctions via the dtruth predicate and a suitable ordinal notation.

Unfortunately, Field's position on the current matter is (well) beyond the scope of this essay, but the basic issue can be (very roughly) sketched as follows. What is remarkable about Field's proposal is that every sentence in the 'hierarchy of operators', and hence any Liar sentence in the given hierarchy, gets characterized by some (semantically significant) predicate or other—in particular, some predicate constructed out of 'stronger truth' (determinacy), negation, and so on. Call such predicates *indeterminacy predicates.* It would seem that one could *conjoin* all such indeterminacy predicates to get one 'big, unified indeterminacy predicate', or at least get the same effect of such 'conjunction' by quantification and dtruth. All of this seems to be eminently intelligible. The trouble, according to Field, is that the apparent intelligibility is merely apparent.

In short, Field [2005b] shows that such a 'conjunction', to the extent that it is achievable, either cannot be (coherently) constructed or, to the extent that it can, won't behave as expected; the reason, in keeping with a thoroughgoing paracomplete account, is that it is 'indeterminate' what the conjuncts are! Any such quantification, aimed at achieving the (alleged) unified indeterminacy predicate, would at best be restricted quantification involving an 'indeterminate' or 'fuzzy' restricting condition.

The sense in which such 'conjuncts' are 'indeterminate', or the restricting condition 'fuzzy', is quite involved and, unfortunately, too involved for this essay. But it is worth noting that Field has a reply to the charge that he fails to achieve the target sense of exhaustive characterization. For a full, detailed discussion, see [Field, 2005b].[98]

4.3 Paranormal

Kripke showed how we can have a consistent (and, hence, non-trivial) dtruth predicate, but not for a language in which all dT-biconditionals hold or for a language in which we can dtruly characterize Liar-like sentences (let alone 'exhaustively' so characterize). Field, in turn, shows how we can have a consistent dtruth predicate in a language that validates all dT-biconditionals and enjoys the resources to characterize all Liars constructible in the language (or, at least, in the 'hierarchy of determinacy operators'). What Field does not—and, in his framework, cannot consistently—do is offer a 'unified predicate' that characterizes all Liar-like sentences. Deviating from Field's framework, this section (briefly) sketches a route one might take towards achieving a unified 'classification' of Liars.[99]

Paracomplete theorists recognize the need to expand our otherwise classical

[98] While Field's proposal is clearly an advance on earlier paracomplete accounts, I should note that questions remain open as to whether Field's proposal sufficiently covers related semantical paradoxes, in particular, paradoxes of denotation (e.g., Berry's). Graham Priest [2005] argues that Field can handle such paradoxes, if at all, only via ad hoc restrictions. Field [2005a] disagrees, maintaining, as any paracomplete theorist will, that Berry's paradox is problematic only if some illicit use of LEM is invoked. But the issue is tricky and, alas, beyond the scope of this essay.

[99] This section is drawn from [Beall, 2006].

categories. We have dtrue sentences and, derivatively, dfalse sentences; and there are 'other' sentences 'in addition to the normal ones', sentences like Liars. Call such sentences *paranormal*, again from the Greek 'para'. How, if at all, can we achieve a consistent dtruth predicate, all dT-biconditionals, and also 'exhaustive characterization' in the target (but, admittedly, imprecise) sense, one for which we have a unified 'paranormal' predicate? I will briefly sketch an answer, one that, at least philosophically, depends on a particular conception of the *role* of (the predicate) 'paranormal'. I will not discuss a suitable conditional, but it should be clear that something along the lines of Field's 'neighborhood' conditional will be available.[100]

Philosophical picture

What the Liar (and its ilk) teaches us is that besides the dtrue sentences and dfalse sentences, there are paranormal sentences. One might now wonder: what is it to be a 'paranormal' sentence?

The suggestion is that we set the question aside. For present purposes, it suffices merely to 'tag' the target sentences (e.g., Liars) as such, namely, as *paranormal*. Ultimately, there may well be no interesting property of *being paranormal*, and accordingly no hope of informative 'analysis' or explication of 'paranormal'. But the term may none the less serve to give us the sort of 'exhaustive characterization' desired, just by giving us a 'logical device' of sorts with which to 'classify' the target sentences. My suggestion is that we resist questions concerning 'the nature' of paranormals, seeing it merely as a tag (a logical category) introduced for the target sentences.

Notice that even at this stage—without giving much more than a 'classifying role' for the device—Liar phenomena already arise.

$\surd$ The ticked sentence in §4.3 is either not dtrue or paranormal.

And such a sentence itself is surely among the very sort for which we introduced the tag 'paranormal', and indeed the usual Liar-reasoning will suggest as much given the relevant version of EC: namely, that every sentence (hence, the ticked one) is dtrue, dfalse (true negation), or paranormal.

The upshot is that *paranormality and dtruth apparently overlap*. And my suggestion is that we simply accept as much. After all, we want a simple, exhaustive characterization, one that is consistent. And we can enjoy as much by acknowledging that some paranormal sentences—just in virtue of the role of 'paranormal' and the basic expressive job of 'dtrue'—turn out to be dtrue.

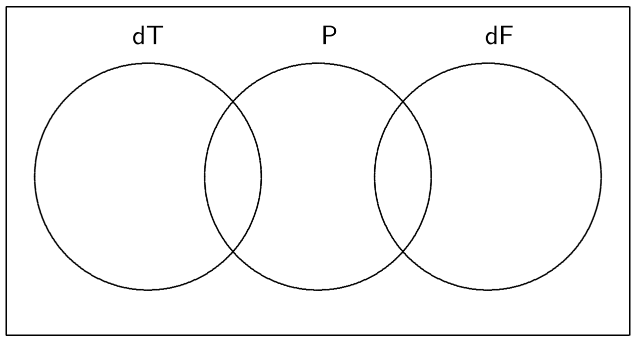
Likewise, of course, various paranormal sentences will inevitably be dfalse, for example,

$\star$ The starred sentence in §4.3 is not paranormal.

[100] Indeed, I'm inclined to think that an even simpler conditional, invoking so-called non-normal worlds, will do the trick. For a sketch of such a conditional—albeit in a paraconsistent setting—see §5.1.

If we have it—via the relevant version of EC—that the starred sentence is dtrue, dfalse, or paranormal, then it's dtrue iff not paranormal (and, so, not dtrue and paranormal). So, the sentence (just reasoning intuitively, at the moment) is paranormal, and so... a dfalse paranormal.

One might picture the story as follows, although one ought to keep in mind that this is only a heuristic.[101]



One might press for analysis or explication: what is it to *be* paranormal?! The suggestion is that we resist the question. Truth itself (at least on a suitably deflationary conception) affords little by way of informative analysis. On the usual picture, we began with our dT-free fragment and had no problems except expressive ones due to our finite limitations. We could neither implicitly nor explicitly assert everything that we wanted to assert. Towards that end, our 'dtruth'-device was introduced. But once 'dtrue' was introduced (into the grammatical environment of English), various unintended sentences emerged—typical Liars and so on. Towards 'classifying' *those* sentences, 'paranormal' is introduced. But given the job of 'paranormal', there's little reason to expect—let alone demand—an informative analysis. Indeed, as is evident, there is even less to say about 'paranormal' than 'dtrue'.[102]

Notice that by allowing 'overlap' between the paranormal and the dtruths, we thereby avoid the need to invoke infinitely many (non-unifiable) 'stronger truth' predicates. Once 'paranormal' is introduced, unintended by-products of it emerge—*this sentence is paranormal*, etc. The suggestion is that we simply let such sentences be among the paranormal, even though—given the role and rules of 'dtrue'—they may likewise be dtrue. (Similarly for 'dfalse'.) If our chief concern is to 'exhaustively characterize' or 'classify' in a *consistent* and *simple* way, then such overlap is harmless, provided that 'dtrue' and 'dfalse' avoid overlap.

Notice, too, that the proposal is *not* motivated by a search for some suitably stronger notion of 'truth' with which we can assert the 'un-truth' (in a stronger

[101] And see §4.3 for an alternative (and, in my view, more attractive) picture.
[102] But see §4.3 for an alternative picture.

sense of 'true') of Liars. The proposal is simply that we classify such sentences as, well, paranormal—and that is that. If more is required for serious inquiry (e.g., science), then the proposal is wanting in that respect. But, as far as I know, not more is required. Moreover, one can, of course, define a stronger notion of truth via dtruth and 'paranormal' if one wishes. (See *Strong truth*, page 228.) But doing so seems not to be a requirement.

There may be philosophical issues to be sorted out, but they must wait for another time. For now, I briefly turn to a simple, formal sketch of the idea.

A formal model

The picture is along familiar many-valued lines.[103] Our 'semantic values' (in the *formal story*) are elements of $\mathcal{V} = \{1, \frac{3}{4}, \frac{1}{2}, \frac{1}{4}, 0\}$, with designated elements in $\mathcal{D} = \{1, \frac{3}{4}\}$.[104]

Our atomics are interpreted via a function ν in the usual way,[105] extended to compounds along the lines seen in Strong Kleene: $\nu(A \wedge B)$ is the minimum of $\nu(A)$ and $\nu(B)$, and $\nu(A \vee B)$ the maximum. (Quantifiers can be treated similarly, as generalised conjunction and disjunction.) Negation is likewise familiar: $\nu(\neg A) = 1 - \nu(A)$. Hence, negation is fixed at $\frac{1}{2}$ but otherwise toggles designated and undesignated values; it is thus 'normal' in the usual (formal) sense.

We assume a special predicate dT to be interpreted as a dtruth predicate: $\nu(\mathsf{dT}\langle A\rangle) = \nu(A)$ for any 'admissible' ν. Falsity, in turn, is derivative: $\nu(\mathsf{dF}\langle A\rangle) = \nu(\mathsf{dT}\langle\neg A\rangle)$.

Finally, we add a unary connective π (our 'paranormal' device), which is interpreted thus:[106]

$$\nu(\pi A) = \begin{cases} 0 & \text{if } \nu(A) \in \{1, 0\} \\ \frac{3}{4} & \text{otherwise.} \end{cases}$$

Note that—letting 'P' be our 'paranormal' predicate—the extension of P, namely,

$$\mathcal{P}^{+} = \left\{A : \frac{1}{4} \leq \nu(A) \leq \frac{3}{4}\right\}$$

may well be negation-inconsistent; it may well contain both A and $\neg A$ for some A. (Indeed, for typical π-free Liars, that will be the case.) In this respect, *being*

[103] NOTE WELL: While I do not (here) discuss a suitable conditional, the conditional in question is *not* to be constructed along standard Łukasiewicz lines (for reasons mentioned in §4.1.). The conditional will be one that is not 'truth-functional' but, rather, either invokes 'non-normal points' familiar from Kripke and, more recently, work in 'relevant' semantics (see §5.1 for brief sketch), or something along the lines of Field's 'neighborhood'.

[104] Logical consequence – semantic validity – is defined as usual in terms of $\mathcal{D}$. I skip the definition here.

[105] Note that if, for some reason, one wanted to have a three-valued language, letting $\mathcal{V}$ contain only three values, one could let interpretations be non-functional relations and – given appropriate constraints on the interpretations – thereby achieve the same framework. Functional interpretations are more familiar, and so I go with that.

[106] Note that I will use 'π' for both the connective and operator, trusting that context will do its clarifying job.

paranormal differs from *dtruth* (similarly, *dfalsity*), as the extension of the latter, namely,

$$\mathcal{T}^{+} = \left\{ A : \frac{3}{4} \leq \nu(A) \right\}$$

is always negation-consistent.[107]

Comments

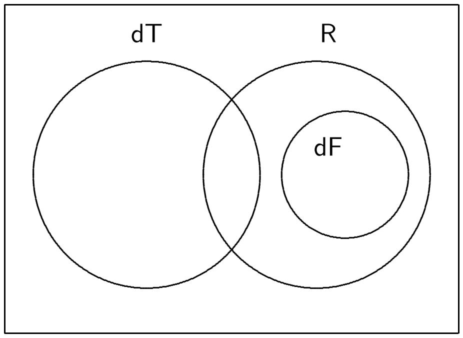
What the 'paranormal' approach yields is a simple, consistent way towards achieving exhaustive characterization in the target sense, one that employs a unified 'paranormal' predicate. Liars, one and all, may safely be classified as paranormal, even though some paranormals may also be dtrue or dfalse. And, in general, we achieve 'unified exhaustive characterization', in that every sentence will fall under one of our three semantically significant predicates: either dtrue, dfalse, or paranormal. In this section I briefly address two issues, leaving further discussion to other work.

An alternative picture: residual

As above, the 'paranormal' account achieves a simple, unified exhaustive characterization (while, suitably filled out, preserving the salient virtues of other paracomplete accounts). Of course, such exhaustion—such 'unification', as it were—comes at the 'cost' of there being nothing illuminating to say about *being paranormal.* How much of a cost that may be is not entirely clear. Still, one might at least like to have it that, for example, if a sentence is not dtrue, then it is paranormal, and if not paranormal, dtrue. We do not have as much on the foregoing account, but one can easily tweak the picture to get as much.

One might, for example, settle on a (fairly common) thought according to which we have our *dtrue* sentences and, beyond that, we simply have *the rest*—be they dfalse or whathaveyou. Here, the picture is one according to which our goal (e.g., in science or rational inquiry, generally) is to record the full, dtrue story of our world and simply chuck the 'remainder' to the side. On this picture, anything *dfalse* (having true negation) is 'residual' and, moreover, anything *not residual* dtrue. The picture might look as follows, with dfalsity now merely a proper part of the residual.

[107] A consistency proof cannot be done in the usual fashion, given that π is non-monotonic, but such a proof is available. (Thanks to Tim Bays and Greg Restall for their interest in the idea; each of them independently suggested different proof-sketches.)



Such an idea requires only a minimal tweak of the 'paranormal' account. In particular, instead of 'paranormal', define our 'residual' tag thus:[108]

$$\nu(\dagger A) = \begin{cases} 0 & \text{if } \nu(A) = 1 \\ \frac{3}{4} & \text{otherwise.} \end{cases}$$

In turn, the predicate 'R' (residual) behaves thus:

$$\mathcal{R}^{+} = \left\{ A : 0 \leq \nu(A) \leq \frac{3}{4} \right\}$$

$\mathcal{R}^{+}$ will differ from 'dfalse' by being (negation-) inconsistent, whereas the extension of 'dfalse' will be (negation-) consistent:

$$\mathcal{F}^{+} = \left\{ A : 0 \leq \nu(A) \leq \frac{1}{4} \right\}$$

This variation reflects the 'residual' conception according to which any dfalse sentence is residual, and any sentence that is not residual is dtrue. The salient difference between π and $\dagger$ is that we have

$$\neg A \Vdash \dagger A$$

and

$$\neg \dagger A \Vdash A$$

only for $\dagger$, not for π.[109] This reflects the idea, as above, that (in some sense) science—or, generally, rational inquiry—aims to separate dtruth from the 'remainder', the residual. But, as before, once we add (even a minimalistically construed) category 'residual', EC will demand overlap, and so we allow that some of the resulting residuals are true. But no matter; the given overlap is harmless, and also allows for a simple 'unified exhaustive characterization'.

[108] To avoid confusion, I will use '$\dagger$' for our target 'residual' tag (rather than simply redefine 'π').

[109] While I am not discussing a suitable conditional here, the (target) conditional versions of the above principles will similarly hold.

Strong truth

Let me make it plain that the foregoing account is not motivated by a desire to say that Liars are in some sense 'not true'. Still, one can say as much if one wants. For example, on the 'paranormal' account (but similarly for 'residual'), define *robustly true* thus:

$$\mathbb{T}A \quad \text{iff} \quad A \wedge \neg\pi A$$

Obviously, we have it that $\nu(\mathbb{T}A) = 1$ iff $\nu(A) = 1$. As above, $\nu(\pi A) \in \{\frac{3}{4}, 0\}$ and, hence, $\nu(\neg\pi A) \in \{\frac{1}{4}, 1\}$ for all ν.

We immediately get principles reminiscent *but atypical* of standard 'determinately' operators.

* $\mathbb{T}A \Vdash A$.
* $A \nVdash \mathbb{T}A$. (Just let $\nu(A) = \frac{3}{4}$.)

The second principle is what is atypical with respect to standard 'determinately' operators, especially since $A \rightarrow \mathbb{T}A$ will likewise fail, at least on the intended construction (about which, for space reasons, I've said little here). While I won't try to defend such 'failures', it seems to me to make sense: we are acknowledging that some paranormals may be dtrue (true in our basic, merely expressive, entirely transparent sense) but not thereby 'robustly true', and that this point carries all the way through—both for the 'validity' reading and the conditional reading.[110]

The advantage (if any) of having such 'robust truth' is that it may do some work that our fundamental dtruth predicate isn't cut out to do. Our fundamental dtruth predicate has only the job of being a transparent generalization device, of being such that $\mathsf{dT}\langle A\rangle$ and A are intersubstitutable for any A, including, of course, the paranormal A. The value (if any) of 'robust truth' is that it may cut distinctions that dtruth itself can't—and was never intended to—cut. In particular, one can dtruly say that, for example, the ticked sentence in §3 is not robustly true. Of course, what, if anything, one gains from this—over and above merely asserting that the given sentence and its negation are 'residual' (or paranormal)—is not obvious. But, for space reasons, I will leave the matter there.[111]

5 Paraconsistent

'Para' here is used as in 'paracomplete' (see §4), and 'consistent' for negation consistency, the idea being that we are moving 'beyond' typical (negation-) con-

[110] That these 'readings' inevitably come apart in paracomplete (and paraconsistent) approaches is briefly discussed in §6.

[111] I will note one more thing. $\mathbb{T}$, defined as above from our 'paranormal' (similarly 'residual') device, seems to behave very much like—if not exactly like—the KF-truth predicate. This is philosophically interesting, as it gives us the (e.g., expressive) virtues of dtruth but likewise a stronger notion of truth that has gained independent support from others. But details are left for elsewhere.

sistency constraints.[112] What is common to paracomplete theories, as in §4, is a rejection of LEM. What is common to *paraconsistent* theories, in turn, is a rejection of EFQ (ex falso quodlibet or, more colorfully, explosion), the rule according to which (arbitrary) B follows from (arbitrary) $A \wedge \neg A$ or from $\{A, \neg A\}$.[113]

A paraconsistent logic, then, is one in which EFQ fails: $A, \neg A \nVdash B$. An *explosive logic* is one in which EFQ holds. Any theory—hence, any dtruth theory—according to which both A and $\neg A$ are dtrue is *trivial* if its underlying logic is explosive. Paraconsistent logics afford non-trivial but (negation-) inconsistent theories. And that's the basic idea behind paraconsistent dtruth theories, in general.

While paraconsistent approaches to truth are far from dominant (to say the least), there have been various proposals either directly in the tradition or very close to the spirit of it, including [Dowden, 1984], [Priest, 1979; Priest, 1987], [Rescher and Brandom, 1979], [Shaw-Kwei, 1954], [Visser, 1984], [Woodruff, 1984], variations of [Yablo, 1993a; Yablo, 1993b], and others.[114] For present purposes, I will focus on Priest's so-called dialetheic account.[115]

I should note that in paraconsistent theories, as in their paracomplete relatives, 'reject A' is not equivalent to 'accept $\neg A$', and probability requires modification (in effect, the dual of the paracomplete modification). For discussion of (minor) adjustments to probability theory, see [Priest, 1987].

5.1 Priest

Graham Priest takes both NTP and ECP seriously, and, in a large body of work, argues that only by acknowledging 'gluts', sentences that are both dtrue and dfalse, do we achieve satisfactory answers to those projects.[116] Priest (and Richard Sylvan, formerly Routley) coined the term 'dialetheism' (die-a-*lee*-thism) for the view that some sentences are both dtrue and dfalse—equivalently (in the given frameworks), that some dtrue sentences have dtrue negations. Whether dialetheism is the only satisfactory approach to NTP and ECP is something that I leave open—or, at least, to (arguments in) Priest's cited work. My aim, as throughout, is only to sketch the basic idea.

[112] Actually, paraconsistentists seem to have different views of the import of 'para', but I will ignore this here. For discussion, see [Priest, 2000].

[113] Some well-known paraconsistent logics are such that 'conjunction' is abnormal—either Simplification or Adjunction fails. See [Priest, 2002] for broad discussion. I will skip discussion of such logics and focus entirely on 'normal' conjunction.

[114] Martin, Dowden, and Yablo all (independently) tweak Kripke's iterative proposal to achieve an 'inconsistency' account. (Also, Field [2005a] briefly discusses a paraconsistent 'dual' of his own iterative construction.) Rescher–Brandom's and Priest's accounts are compatible with such an iterative construction, but they instead give a simple, non-iterative account. Visser's given discussion is an excellent discussion of fixed point technicalities in four-valued frameworks.

[115] Priest [1987] does not actually propose an account of *dtruth*, and instead imposes various restrictions on truth to avoid full intersubstitutivity. None the less, I will present the basic idea in terms of dtruth.

[116] Priest uses the terms 'dialetheia' and 'true contradictions' for *gluts*. I will use 'gluts', which is a term coined by Kit Fine [1975].

I should note one caveat: that *dialetheism* and *paraconsistency* are not one and the same! While all paraconsistent logicians see paraconsistent logic as a useful tool for modeling inconsistent but non-trivial theories (e.g., naïve truth, naïve semantical properties, etc.), many—if not most—reject that such theories are 'possibly dtrue'. Dialetheists are a small minority, holding that some of the given theories are not only possibly dtrue, but actually so.[117]

Philosophical picture

According to dialetheism, what the Liar paradox teaches us is that some dtruths have dtrue negations, specifically, Liars.[118] The lesson, I think, is quite natural. After all, 'dtrue', as the (typical) story goes, was introduced solely as an expressive, entirely transparent device: for *any* sentence A, be it dT-ful or otherwise, $\mathsf{dT}\langle A\rangle$ and A are intersubstitutable. But, of course, once the device (a unary predicate) is introduced into the language, Liar-like sentences emerge, sentences that say of themselves only that (for example) they are not dtrue, that they are dfalse. Given LEM, such sentences turn out to be dtrue and dfalse.

The *paracomplete* response, as above, is to reject LEM, preserving the (negation) consistency of dtruth. The *dialetheic* response is to accept the inconsistency of dtruth but reject EFQ, thereby preserving the non-triviality of dtruth. Both approaches agree that our 'classical categories' need to be expanded—there's something 'in addition to' the normal categories. The dialetheist's additional category is *glutty*, that is, *both dtrue and dfalse*, a category characterized (in the language itself) just so, 'is dtrue and dfalse'.[119]

Notice that rejecting EFQ is not ad hoc. To begin, why think that (arbitrary) B follows from (arbitrary) A and $\neg A$? One answer might be that there's no apparent counterexample. Indeed, one (many?) might think it 'inconceivable' that both A and $\neg A$ be dtrue, and hence that, vacuously, the given inference count as valid. Such an answer, I think, is entirely natural and, in general, quite reasonable; however, it is not ultimately sufficient. On one hand, intuitions are 'built' from normal, run-of-the-mill cases—usually cases 'grounded' in non-semantic facts (in a minimal sense of 'facts')—and, in such cases, it is difficult (to say the least) to understand what it would be like for $A \wedge \neg A$ to hold.[120] On the other hand, EFQ is a principle about *all* cases (all sentences), not just the run-of-the-mills. When

[117] For discussion of different 'grades of paraconsistency', see [Priest, 2000]. (And for a reply to Priest's alleged 'slippery slope' towards dialetheism, see [Beall and Restall, 2005, Part IV].)

[118] Note that whether a given Curry sentence is a Liar depends on the conditional. If the conditional is the hook (material conditional), then dialetheists (at least those, like Priest, who accept LEM) will count the sentence among the gluts. (Such a sentence will be nothing but a familiar disjunctive Liar 'I am not dtrue or everything is dtrue'.) If a genuine (detachable) conditional is at work, then dialetheists—on pain of triviality—will reject the sentence. See §5.1 for brief discussion of a suitable conditional.

[119] I will briefly address some 'characterization' worries in §5.1.

[120] But, alas, Priest [1987] argues that even in non-semantic, sometimes run-of-the-mill cases, there is good reason to accept $A \wedge \neg A$ (for some A). I will ignore such considerations here, for space reasons.

one takes Liars into consideration, especially with respect to (entirely transparent) dtruth, it is fairly easy to see how for some A, namely, an A that says of itself (only) that it is dfalse, both A and $\neg A$ could be dtrue. (Just consider the intuitively correct reasoning of the Liar paradox!)

Another reason one might give for the validity of EFQ invokes 'the very meaning of negation'. One might say that what we *mean* by negation is such that EFQ holds. But this reply either begs the question or is confused. After all, part of the job of a logical theory is to 'describe' the behavior of our various so-called logical connectives, including negation. The question at hand is whether EFQ holds. To say that EFQ holds in virtue of the very meaning of negation is to either beg the question or confuse theory with subject matter.[121] Given the role of dtruth and the existence of Liars, it is not obviously unreasonable to conclude, as the dialetheist does, that some sentences are both dtrue and dfalse, and hence that some A is such that $A \wedge \neg A$ is dtrue.

That rejecting EFQ is not ad hoc is now plain. If one takes seriously the idea that a sentence—say, a Liar—could be both dtrue and dfalse, then arguments for the validity of EFQ are hard to find. C. I. Lewis' famous 'independent argument' for EFQ [1932] is faulty precisely for its not taking the initial supposition seriously. After all, (seriously) suppose that A and $\neg A$ are (both) dtrue, in which case, presumably, $A \vee B$ is dtrue, since, as above, at least one of the disjuncts is dtrue (viz., A). But now the point: there is no reason at all to think that B follows from the (supposed) dtruth of $\neg A$ and $A \vee B$. Indeed, it is easy to see why this inference—namely, Disjunctive Syllogism—would fail, at least if both A and $\neg A$ could be dtrue.[122]

What about ECP? Does the dialetheist face the predicament of paracomplete theorists (e.g., Field) who, towards characterizing all Liar-like sentences, are forced to acknowledge 'stronger and stronger truth'? No, at least not obviously. (But see §5.1 for some muted worries.) The paracomplete theorist's trouble emerges from rejecting LEM, and thereby losing the means by which to classify Liars (for which LEM fails). (This is particularly salient in Kripke's proposal, as above.) But the dialetheist need not reject LEM. Indeed, Priest [1987] argues that once EFQ is rejected (on the basis of recognizing gluts), there is no good argument for rejecting LEM.[123] Moreover, and perhaps more centrally, the dialetheist essentially recognizes overlap among her semantically significant categories. Such overlap, as in the 'paranormal' approach, affords a unified predicate in terms of which all Liars may be classified.

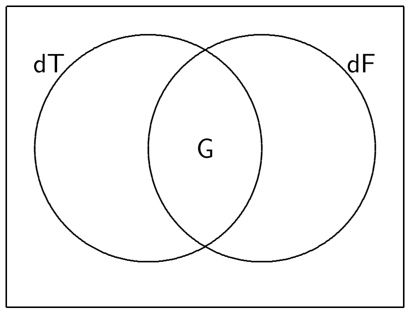
[121] Quine's famous 'dilemma of the deviant logician' seems to make just such a confusion. See [Quine, 1970].

[122] For more on 'the very idea of gluts', and related issues, see [Priest, 1998].

[123] I will briefly return to this issue in §5.2, wherein a 'gappy and glutty' (dialetheic) position is (very briefly) suggested.

Formal picture

Priest's basic picture, as above, is one according to which our *Other* category—*glutty*—overlaps both dtrue and dfalse; indeed, the additional category just is the given overlap (with G for *glutty*). A picture is as follows.



This picture is the 'dual' of the Strong Kleene picture, which likewise has three semantically significant categories, but, unlike the above, none of the Strong Kleene categories overlap. Indeed, Priest's purely extensional LP [1979] just is the dual of K_3, and so clauses for conjunction (minimum), disjunction (maximum), and negation remain exactly as before, namely (K1)–(K3). (See §4.1.)

LP, then, is exactly like K_3 except that the 'middle value' is designated. Designating the middle value, which represents 'both dtrue and dfalse', reflects the idea that any dtrue sentence—even if also dfalse—is assertible. Letting $\mathcal{T}^+$ and $\mathcal{T}^-$ be the extension and antiextension of dT, and letting $\mathcal{S}$ comprise all sentences (in the augmented dT-ful language), the chief contrast between K_3 and LP, at least semantically, comes to this:

K3. Every interpretation is such that $\mathcal{T}^+ \cap \mathcal{T}^- = \emptyset$ but some interpretations are such that $\mathcal{T}^+ \cup \mathcal{T}^- \neq \mathcal{S}$.

LP. Every interpretation is such that $\mathcal{T}^+ \cup \mathcal{T}^- = \mathcal{S}$ but some interpretations are such that $\mathcal{T}^+ \cap \mathcal{T}^- \neq \emptyset$.

So, while paracomplete approaches put Liars 'beyond' $\mathcal{T}^+ \cup \mathcal{T}^-$, Priest's (dialetheic) paraconsistent approach puts Liars into $\mathcal{T}^+ \cap \mathcal{T}^-$.

Validity is defined exactly as in the K_3 framework: $\Sigma \Vdash A$ iff every interpretation that designates (each element of) Σ also designates A. In other words, $\Sigma \Vdash A$ is valid iff for all interpretations ν and all $B \in \Sigma$, if $\nu(B) \in \{1, \frac{1}{2}\}$, then $\nu(A) \in \{1, \frac{1}{2}\}$.

With validity so defined, it is immediately clear that LP is paraconsistent, that is, that $A, \neg A \nVdash B$. Just consider an interpretation such that $\nu(A) = \frac{1}{2} = \nu(\neg A)$ but $\nu(B) = 0$.

Similarly, the above counterexample to EFQ also shows that disjunctive syllogism (DS) is invalid in LP: $A \vee B, \neg A \nVdash B$. This 'failure' is especially significant inasmuch as DS is equivalent to 'material modus ponens', the rule that B follows from $A \supset B$ and A, where $A \supset B$, as before, is defined $\neg A \vee B$. But, then, LP fails to have a genuine conditional, an issue to which I return in §5.1.

So goes the basic (extensional) framework. The resulting account is one that answers NTP by rejecting EFQ, and answers ECP by recognizing a unified predicate (say, 'glutty') applying to all Liar-like sentences. I turn to a few issues, including the issue of a suitable conditional.

Comments

In this section I address only a few philosophical issues that arise with respect to Priest's dialetheic approach to dtruth.[124]

Non-triviality project

Recall that NTP is the project of explaining how we can enjoy a non-trivial language that has a dtruth predicate and Liar-like sentences. In Priest's case, the project is not to show how we can enjoy a *consistent* dtruth predicate despite Liars, but how we can enjoy non-triviality. As in §4.1, one generally aims to answer NTP by constructing an artificial, formal language—the *model language*—that contains its own dtruth predicate, and then claims that, at least in relevant respects, 'real dtruth' is modeled by truth-in-the-model language.

Just as in the Kripke case, one might object that Priest's account doesn't sufficiently answer NTP, the reason being that certain notions used in the 'metalanguage' are not expressible in the object- or 'model language'. In particular, one might argue just as in the Kripke case: an LP-based account has it that some sentences are 'both true and false', some 'simply true', and some 'simply false'. Let $\mathcal{L}_m$ be a typical LP-based 'model-language'. In a classical metalanguage, one defines '$\mathcal{L}_m$-simply true', '$\mathcal{L}_m$-simply false', and, of course, '$\mathcal{L}_m$-both-true-and-false'. Since such notions are defined via classical set theory, one can, in turn—sticking within the (classical) metalanguage—prove that $\mathcal{L}_m$ cannot define, for example, $\{B : B \text{ is } \mathcal{L}_m\text{-simply false}\}$. The charge, then, is that, on pain of triviality (versus 'mere negation-inconsistency'), $\mathcal{L}_m$ fails to be an adequate model of real dtruth, since $\mathcal{L}_m$-truth achieves non-triviality in virtue of lacking (the given) notions that we have expressed in our real language.

So put, the charge faces the same problem as the related Kripke case: it is either confused or unwarranted. First, notice that classical logic is an extension of the logic LP. Semantically, every classical interpretation is an LP interpretation, and thus Priest's model language can enjoy a proper, classical fragment. The idea, of course, is that we are classically modeling our 'real, non-classical language'. But, then, notions defined squarely within the classical metalanguage—a proper fragment of one's real language—are merely model-relative ones that, not surprisingly,

[124] Actually, once again, I should note that Priest doesn't himself give an account of *dtruth*, but his framework, as above, obviously affords such an account. See [Priest, 1987] for Priest's preferred account of truth.

behave entirely classically. What one needs to show is that there are non-model-relative notions that are expressed in our real language but not expressible in the model language. Pending some (as yet not given) reason to think that the real language has some non-model-relative 'untruth predicate' that behaves over our whole (real) language as our model-relative '$\mathcal{L}_m$-simply false' behaves over $\mathcal{L}_m$, the given charge is either unwarranted or simply curious, given the aim of $\mathcal{L}_m$.

Exhaustive characterization?

One might think that the dialetheist does not achieve exhaustive characterization in the target sense.[125] After all, the dialetheist maintains that all sentences are either *just dtrue*, *just dfalse*, or *both dtrue and dfalse*. But what predicate does the dialetheist use to express 'just dtrue'? In the formal story, 'just dtrue' is modeled by 1, 'just dfalse' by 0, and 'both' modeled by $\frac{1}{2}$. But, again, how is 'just dtrue' or 'just dfalse' expressed in the real language? If 'just dtrue' is a semantically significant category *distinct from* 'dtrue', then the dialetheist needs an account of it. Unfortunately, no obvious candidate emerges. (One might think of something along the lines of $\neg(A \wedge \neg A)$, but this is valid in LP, and hence cuts no distinctions.)

Another way putting the worry is as follows. In order to characterize simple dT-ful Liars, standard paracomplete theorists invoke some stronger notion of truth, say $\mathbb{T}$. But, now, in order to characterize $\mathbb{T}$-ful Liars, such theorists invoke an even stronger notion of truth, say $\mathbb{T}'$. And so on. But now consider Priest's dialetheic approach. While there is no problem characterizing Liars—they are both dtrue and dfalse—there is an apparent problem characterizing the 'normal sentences', the 'non-dialetheia', the sentences that are 'just dtrue' or 'just dfalse'. The problem seems to be exactly analogous to the standard paracomplete theorist's problem. Whereas the standard paracomplete theorist is pushed towards a stratified conception of 'strong truth', the dialetheist, at least on the surface, seems to be pushed towards a stratified conception of 'just dtrue'.[126]

I am not sure what dialetheists should say about this (admittedly, as yet vaguely sketched) worry. Priest [1987] has long maintained that 'just dtrue' is as inconsistent as dtruth itself, a claim that is neither unreasonable nor surprising. (Consider a sentence like 'this sentence is just dfalse' or the like.) But such a response does not obviously get to the heart of the worry. The point of rejecting EFQ is to allow *some* sentences to be *both dtrue and dfalse* while still enjoying (many) sentences that are 'just dtrue'. This central claim of dialetheism utilizes what one would take to be dialetheism's key semantic categories—*both dtrue and dfalse*, and *just dtrue*. But if that is right, and if 'just dtrue' is supposed to be distinct from 'dtrue' (as it is in the formal model, where only the latter is represented by all designated

[125] The following issue is related to discussion in both [Parsons, 1990] and [Shapiro, 2004].

[126] I should note too that in Field's framework, just such a stratified notion of 'just dtrue' is available to the dialetheist. (Of course, one would—in effect—dualise the Field construction.) In particular, as α increases, $\mathbb{D}^{\alpha}A$ rules out more and more of the gluts, the 'dialetheia'. See [Field, 2005a] for some discussion.

values), then the dialetheist doesn't obviously achieve full characterization in the target sense. But I will leave the matter there.[127]

A suitable conditional

With respect to a conditional, Priest's proposal notably contrasts with the K_3 Kripkean proposal (§4.1). While Kripke's proposal enjoys a genuine (i.e., detachable) conditional, it fails to enjoy a conditional that validates all dT-biconditionals. Priest's LP, on the other hand, enjoys a conditional that validates all dT-biconditionals, since $\Vdash A \supset A$ in LP; however, the conditional isn't genuine, since it fails to detach.

Unlike Kripke, Priest (like Field) takes the issue of a suitable conditional seriously, where a 'suitable conditional' is a *genuine* conditional that avoids Curry (e.g., Contraction, etc.) and validates all dT-biconditionals. While Priest has suggested various accounts, I will briefly indicate a particularly simple one [Priest, 1992].[128]

Since the basic issue concerns Curry and detachment, I will here ignore negation. (As will be evident, the framework affords various approaches to negation, including a natural LP-approach.) We expand the language along modal lines, invoking points of evaluation—worlds. Exactly how this is done is not pressing, for present purposes. To make matters simple, we will take a 'propositional' approach, letting $|\,|$ be a function from atomic sentences into $\wp(\mathcal{W})$, the idea being that $|A|$ is the 'proposition' expressed by A, the set of worlds at which A is true. Our set of points $\mathcal{W}$ is the union of two sets, $\mathcal{N}$ (normal points) and $\mathcal{NN}$ (non-normal points),[129] with a distinguished element $@ \in \mathcal{N}$ (the actual world) and $\mathcal{N} \cap \mathcal{NN} = \emptyset$. With respect to extensional connectives (here ignoring negation), $|\,|$ is expanded as one would expect.

$$|A \wedge B| = |A| \cap |B|$$

and

$$|A \vee B| = |A| \cup |B|$$

Finally, interpretations come equipped with an 'arbitrary evaluator' ϑ the task of which is to assign values to $\rightarrow$-claims at non-normal points: ϑ is a function

[127] A related—but, in my opinion, not terribly troubling—issue is that, for example, an LP-based approach to dtruth is committed to saying of Liars that they are neither dtrue nor dfalse; for, such LP-based dialetheists assert $A \wedge \neg A$ for any Liar A, but that is equivalent in LP to $\neg(A \vee \neg A)$, which, given intersubstitutivity, would amount to the claim that A is neither dtrue nor dfalse. See [Field, 2005a] for discussion.

[128] For Priest's latest thoughts on a suitable conditional, see [Priest, 2006a; Priest, 2006b]. For a variation of the approach given below, see [Beall, 2005b], which is closely related to the proposal in [Beall *et al.*, 2005].

[129] Non-normal points were first invoked by Kripke [1965] to model Lewis systems weaker than S4 (systems in which Necessitation fails). [Routley *et al.*, 1982] and [Routley and Loparic, 1978] invoke such points for purposes closer to the current project, as does [Mares, 2004].

from pairs of 'propositions' into $\wp(\mathcal{NN})$. Assuming standard S5 conditions on $\mathcal{W}$, conditionals are evaluated thus,

$$|A \to B| = N \cup NN$$

where N, comprising the *normal* worlds at which $A \to B$ is true, is such that

$$N = \mathcal{W} \text{ if } |A| \subseteq |B|, \text{ and otherwise } N = \emptyset$$

and NN, the *non-normal* worlds at which $A \to B$ is true, is given by ϑ,

$$NN = \vartheta(|A|, |B|)$$

The idea, put (perhaps) more simply, is just this: in addition to 'normal worlds' (among which is the actual), we also recognize *non-normal worlds.* Non-normal worlds are relevant only to the (non-extensional) conditional; extensional connectives behave normally. Conditionals are evaluated differently—indeed, entirely arbitrarily—at non-normal worlds. In effect, conditionals are evaluated exactly as you would expect at 'normal worlds', but evaluated in any manner one likes at non-normal worlds (provided the given extensional clauses are respected). This would wreak havoc if *validity* were defined over *all* worlds (of all interpretations), but it needn't be, and in fact is not.

Validity is defined in terms of *actual verification* (See §4.2. Note that one could—and standardly does—say that validity is 'truth-preservation' over all normal worlds, but restricting to just @ makes no difference.) In the present context: A is *actually verified* in an interpretation iff $@ \in |A|$.) So long as there's no interpretation such that $@ \in |A|$ but $@ \notin |B|$, then $A \Vdash B$. Similarly, valid sentences are those that are actually verified on all interpretations.[130]

This simple framework might not yield everything that one wants from a conditional, but, for present purposes, it yields the target desiderata. For example, the conditional is 'genuine', since no interpretation actually verifies $A \to B$ and A without thereby actually verifying B. Moreover, $A \to A$ is valid (as brief reflection indicates), and hence all dT-biconditionals are verified. Furthermore, the troubling contraction principles fail, thanks to non-normal worlds. For example, just consider an interpretation according to which $@ \in |A|$ and $|A| \subseteq |B|$, but such that $w \in \vartheta(|A|, |B|)$ but $w \notin |B|$. This serves to invalidate $A \wedge (A \to B) \to B$, and similar counterexamples invalidate the other Curry-generating principles.

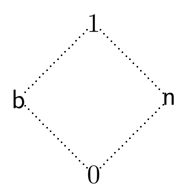
There are other approaches that one might take to conditionals in a paraconsistent context, including, of course, tweaking Field's 'neighborhood' approach. For now, the point is simply that there are suitable conditionals available to the dialetheist.

[130] Of course, there are broader notions of validity that can be defined, but the given one is most relevant here.

5.2 Gaps and Gluts?

Priest's dialetheic proposal offers a natural and very simple approach towards dtruth and, in particular, NTP and ECP. On the other hand, one might be open to such 'gluts' but none the less think that there are also *gaps*, sentences that, somehow, are dtruly classified as 'neither dtrue nor dfalse'.[131] Of course, as in §4.1, it makes little sense, and at any rate is incorrect, to say that some sentences are 'neither dtrue nor dfalse', at least given (only) a Strong Kleene account of negation. A natural response to the problem—not radically different from the *strong truth* tradition—is to acknowledge another negation, a more robust (but logically weaker) *gap-closing* negation, in particular, a sort of 'exclusion negation'. The reason that this is not normally pursued is that such a negation—one for which LEM holds—inevitably gives rise to inconsistency. But this is no major worry if the overall framework is paraconsistent: we may accept LEM for one of the negations (viz., exclusion) but reject EFQ for all negations.[132]

The idea can be modeled using a four-valued language along the lines of Anderson and Belnap's FDE [1975; 1992].[133] Our semantic values $\mathcal{V} = \{1, \mathsf{b}, \mathsf{n}, 0\}$ are ordered thus:



Intuitively, as in Priest's proposal, 1 models sentences that are dtrue but not dfalse, 0 sentences that are dfalse but not dtrue, b sentences that are both dtrue and dfalse, and n sentences that are neither. The designated values $\mathcal{D}$ are 1 and b, the idea being, just as in Priest's proposal, that dtrue sentences are designated (even when they are also dfalse).

Interpretations are functions ν from sentences into $\mathcal{V}$ such that $\nu(A \wedge B)$ and $\nu(A \vee B)$ are the infimum (glb) and supremum (lub) of $\nu(A)$ and $\nu(B)$, respectively.[134]

In FDE we have only (what I shall call) 'choice negation' $\neg$, which toggles 1 and 0 and is fixed at both b and n. We add another negation, pseudo-exclusion $-$, which toggles 1 and 0, is fixed at b, but takes n (gaps) to 1. The result is (what I shall

[131] The material in this section is largely from [Beall, 2005b].

[132] As indicated, paracomplete and paraconsistent theorists unite in rejecting any 'absolute exclusion device', a device $\heartsuit$ such that $\heartsuit A, A \Vdash B$ and $\Vdash \heartsuit A \vee A$. The current suggestion is no different.

[133] The name 'FDE' is now common for the following framework; however, it is perhaps unfortunately so named, since there are various accounts of 'first degree entailment'. But I shall follow what now seems to be common practice.

[134] For present purposes I lay out the propositional semantics; the predicate extension—including the resulting dtruth-theory—is straightforward: one simply allows for both $\mathcal{T}^+ \cap \mathcal{T}^- \neq \emptyset$ and $\mathcal{T}^+ \cup \mathcal{T}^- \neq \mathcal{S}$. (See §5.1.)

call) FDE*.[135] Accordingly, FDE*-interpretations 'obey' the following diagrams with respect to negation.

$\neg$	A
1	0
n	n
b	b
0	1

$-$	A
1	0
1	n
b	b
0	1

Notice that dfalsity, following standard thinking, remains dtruth of negation—dtruth of *choice* negation, as opposed to pseudo-exclusion (henceforth, exclusion).

A *model* of A is an FDE*-interpretation that designates A, that is, an interpretation ν such that $\nu(A) \in \mathcal{D}$. And a model of $\Gamma = \{A_1, \ldots, A_n\}$ is a model of A_i, for each $1 \leqslant i \leqslant n$. Consequence $\Vdash$ is defined thus: $\Gamma \Vdash A$ iff every model of Γ is a model of A. Valid sentences are consequences of $\emptyset$.

With respect to Liars, there will now be two sorts: choice and exclusion. One could treat both sorts of Liar the same, namely, as gluts. On the other hand, one might follow the methodological principle according to which a sentence is gappy if its truth is determined by neither the world (as it were) nor the language. On such a route, choice Liars—similarly, dtruth-tellers (e.g., 'this sentence is dtrue')—are treated as gaps, exclusion Liars as gluts. For more on the philosophical picture, see [Beall, 2005b].

5.3 Comments

It is obvious that the foregoing approach has all the virtues of Priest's dialetheism, at least with respect to NTP and ECP (e.g., unified semantically significant predicates, etc.). Moreover, the conditional indicated in §5.1 may be used for purposes of a suitable conditional—one that detaches and validates the dT-biconditionals. Here, I touch on two issues. Further discussion is available in [Beall, 2005b].

De Morgan and the notion of gaps

As expected, excluded middle fails for choice negation but holds for exclusion: $\nVdash A \vee \neg A$ but $\Vdash A \vee -A$. Moreover, both negations exhibit standard double-negation behaviour, at least in terms of 'inferences'. For example: $A \dashv\Vdash \neg\neg A$ and $A \dashv\Vdash - - A$.[136]

Standard de Morgan laws hold for choice: $\neg(A \vee B)$ is equivalent to $\neg A \wedge \neg B$ (and similarly for the other laws). But exclusion is different: de Morgan laws will

[135] This is not the best name, as it might suggest an approach to FDE using the Routley star, but I trust that no confusion will ensue.

[136] Note, however, that in the double-exclusion case, this is only bi-consequence, not equivalence in the strong sense of 'same value' (which does hold in the choice case). What we have in the exclusion case is co-designation: A and $- - A$ are both designated or both undesignated on any $FDE^\star$-interpretation.

generally hold in one direction but not both. Of particular importance—given the role of exclusion in the notion of *gaps*—is that we have

$$-(A \vee B) \Vdash -A \wedge -B$$

but we do not have equivalence; in fact,

$$-A \wedge -B \nVdash -(A \vee B)$$

A counterexample: $\nu(A) = \mathsf{n}$ and $\nu(B) = \mathsf{b}$. In that case, $\nu(-A) = 1$ and $\nu(-B) = \mathsf{b}$, and so $\nu(-A \wedge -B) = \mathsf{b}$. But, then, $\nu(A \vee B) = 1$, and so $\nu\left(-(A \vee B)\right) = 0$.[137]

Is the 'non-standard' behavior of exclusion—failure of de Morgan principles—a prima facie problem? I see no reason to think as much, in general. Presumably, choice is our 'default' negation; we employ exclusion when we need to talk about failures of choice. Our 'intuitions' about de Morgan, in turn, are presumably based on choice—or, at least, based on 'normal cases', 'settled cases', and so on. That some such (de Morgan) principles should fail for exclusion seems, as said, not to be a problem, in general.

On the other hand, one might worry that the given de Morgan 'failures' pose a problem for the role of exclusion in the notion of gaps. Gappy sentences are supposed to be neither dtrue nor dfalse. But that, one would think, ought to be *equivalent* to saying that such sentences are (exclusion-) not dtrue and (exclusion-) not dfalse. The worry is that such equivalence fails, given that, as above, $-(A \vee B)$ and $-A \wedge -B$ aren't equivalent, in general.

Fortunately, the worry isn't serious: $-(A \vee B)$ and $-A \wedge -B$ *are* equivalent in the special case where B is $\neg A$, which is precisely the case involved in saying that A is neither dtrue nor dfalse. Accordingly, the general failure of de Morgan (for exclusion) seems not to be a particular problem for the notion of gaps.

Strong truth

If, for some reason, one is inclined to acknowledge a notion of 'strong truth', such a notion is definable in terms of our two negations and dtruth [Beall, 2002]. In particular, define a 'robustly true' device $\mathbb{T}$ thus: $\mathbb{T}A$ iff $\mathsf{dT}\langle \neg - A\rangle$, or, equivalently (given dtruth), $\mathbb{T}A$ iff $\neg - A$.[138]

6 Dtruth, validity, and truth-preservation

The notion of validity is often cashed out, at least intuitively, as 'necessary truth-preservation'. At the very least, 'truth preservation' is commonly thought to be a necessary condition of validity. But in the context of *dtruth*, the common

[137] The given 'inference' holds if our values are linearly ordered thus: $1 \succ \mathsf{b} \succ \mathsf{n} \succ 0$. But in that case, de Morgan will break down for choice.

[138] Compare Field's approach (§4.2).

connection is strained, at best. In this section, I briefly discuss a (relatively under-discussed) issue confronting the two common approaches to dtruth—paracomplete and paraconsistent.

One can see the issue in light of the classical picture, wherein we can retain the familiar connection between validity and 'necessary truth-preservation' by giving up a dtruth predicate (for the language) in which to express (in the language) such a connection. (One must resort to a 'richer metalanguage' and, hence, use something other than dtruth.) By contrast, as the foregoing sections have indicated, paracomplete and paraconsistent approaches give us a dtruth predicate; however, the familiar connection between validity and dtruth-preservation is lost (or, at least, strained). I will briefly elaborate. In what follows, I will assume that our basic (extensional) framework is either K_3 (paracomplete) or LP (paraconsistent).

6.1 Validity

Validity, as above, is often thought of as *necessary truth-preservation*, or at the very least *requiring* such truth-preservation. A natural way to understand such a thought is as follows, where Val is a binary predicate having the intuitive sense of *follows from*.

V0. $\mathsf{Val}(\langle A\rangle, \langle B\rangle) \rightarrow \Box(\mathsf{T}\langle A\rangle \rightarrow \mathsf{T}\langle B\rangle)$

where $\rightarrow$ is a genuine conditional, one that at least detaches (in 'rule form').[139] In the case of *dtruth*, which is currently the focus, V0 amounts to

V1. $\mathsf{Val}(\langle A\rangle, \langle B\rangle) \rightarrow \Box(A \rightarrow B)$

Intuitive as V1 may be, problems arise in both paracomplete and paraconsistent settings, at least if V1 is supposed to be expressible in one's given language. One problem, arising from Curry, arises for both approaches.

6.2 Curry

To simplify matters, concentrate just on the weaker principle:

V2. $\mathsf{Val}(\langle A\rangle, \langle B\rangle) \rightarrow (A \rightarrow B)$

As above, we're assuming that $\rightarrow$ is detachable; hence, given plausible assumptions about conjunction, we have

(1) $\mathsf{Val}(\langle A \wedge (A \rightarrow B)\rangle, \langle B\rangle)$

But, then, by V2 and detachment we get

[139] In the present context, speaking of the 'rule form' of Modus Ponens (or any other argument-form), where this invokes a turnstile, is slightly delicate. After all, a turnstile is typically used in a metalanguage (for a formal object language) to represent what, intuitively, we take to be our validity predicate, which is precisely the issue—viz., how to understand a validity predicate *for* our language *in* our language.

(2) $A \wedge (A \rightarrow B) \rightarrow B$

The trouble is that (2) is a notoriously easy recipe for Curry, which results in triviality. The upshot is that, 'intuitively right' as they may appear, V2 (and, hence, V1) must go.

One not altogether unattractive response to this problem is to reject the dominant idea that validity requires 'necessary dtruth-preservation' or, more simply, even dtruth-preservation, and indeed reject that 'validity' be defined at all. Perhaps, instead, one ought to accept that *validity* is a *primitive* notion. And perhaps not all is lost: one might be able to retain various 'intuitive' thoughts about validity but reformulate them in terms of 'acceptance' and 'rejection', notions that are already essential to paracomplete and paraconsistent frameworks.[140]

On the other hand, it would be nicer (in some sense) to retain the idea of 'truth-preservation' if possible. I will consider the paraconsistent and paracomplete cases in turn.

6.3 Paraconsistent: dtruth-preservation

V2 and, hence, V1 must go, as per §6.2. The question is: how, if at all, can we retain the idea that validity requires (let alone *is*) dtruth-preservation?

One route, at least in the LP setting, is to give a 'mixed' account of truth-preservation, one that uses some other 'conditional' in addition to one's genuine one. A natural go invokes $\supset$, the material conditional. For simplicity, we may concentrate on the simpler case:

V2*. $\mathsf{Val}(\langle A\rangle, \langle B\rangle) \rightarrow (A \supset B)$

In an LP-setting, V2* will be dtrue—assuming, as I am, various natural ways of handling $\rightarrow$ (e.g., §5.1). Moreover, Curry paradox is harnessed. We have it that $\rightarrow$ detaches, and so have it that (1) is dtrue. But from V2*, unlike the case of V2, we do not get (2); we get only

(3) $A \wedge (A \rightarrow B) \supset B$

And since $\supset$ does not detach in the LP-setting (and, so, isn't a 'genuine conditional'), the threat of triviality is avoided.[141]

[140] Example: *One ought to reject the validity of a given argument if one accepts that it's possible for the given premises to be true but conclusion false.* This won't generalize to V1 or even V2, at least in the paracomplete framework. (Speaking only of dfalsity will also not quite be enough in the paracomplete case, but I'm simplifying.) Still, in particular cases, such formulations may work. Field (in conversation) has expressed sympathy (if not full endorsement) of this approach.

[141] Let A be a Curry-sentence of the form $A \rightarrow \perp$. Since the dT-schema is underwritten with $\rightarrow$ (our 'genuine conditional'), (3) will yield $\neg A \vee \perp$, which, in LP, is dtrue exactly if A is dfalse, that is, exactly if $A \rightarrow \perp$ is dfalse. Provided that sentences of the form $A \rightarrow B$ may be dfalse (at a point) without A's being dtrue (at that point), the threat of Curry-generated triviality is avoided.

One salient cost of retaining dtruth-preservation via V2* concerns the familiar relation between inference (or, indeed, deductions) and validity.[142] The utility of (knowing that you have) a valid argument is often supposed to be that it is (necessarily) dtruth-preserving. The utility of such dtruth-preservation, in turn, is often supposed to be that our inferences, if dtruth-preserving, 'won't go wrong'. But while V2* ensures that valid arguments are dtruth-preserving (indeed, even necessarily so in LP), the utility of such dtruth-preservation is now dubious: one can know that an argument is valid, that its premises are all dtrue, but none the less be without a valid argument that takes one from such information and the (V2*-) dtruth-preservation of the argument to the given conclusion. In short, the dtruth-preservation of valid arguments is now without any of its familiar pragmatic utility.

To what extent such a cost is crippling remains to be seen. There are (obviously) other ways of 'mixing conditionals' to preserve the idea that validity requires (necessary) dtruth-preservation, but, at least in LP-like settings, most of the options will exhibit the above 'break from pragmatic import'.[143]

6.4 Paracomplete: (robust) truth-preservation

Just as in the paraconsistent setting, the paracomplete account must give up V2 (and, in turn, V1). The question remains: is there any way to preserve that validity requires dtruth-preservation in a paracomplete setting? The answer is No.

The most promising route is something along the lines of V2*, but that will not work. In LP, LEM holds; in K_3, not. Without LEM, V2*, regardless of its other costs, is unavailable.[144] Consider, for example, an instance of V2* in which A is the conjunction of a Liar and its negation, and B is '$1 = 0$'. This immediately shows the inadequacy of V2* in a K_3 setting (wherein EFQ is supposed to be valid).[145]

The upshot is that paracomplete theorists, unlike paraconsistent theorists, cannot enjoy a 'validity' predicate (in the language, for the language) that requires dtruth-preservation. How, then, if at all, can a paracomplete theorist enjoy a 'validity' predicate (for her language, in her language) that, in some sense, is *truth-preserving* (versus dtruth)?

The obvious idea is to invoke 'robust truth' or 'strong truth' or the like. After

[142] Another potential cost concerns the relation between 'counterexamples' (intuitively understood) and validity, but I will ignore this here.

[143] I think that the awkwardness of the given problem can be diminished by drawing a clear distinction between validity (or implication) and inference, much along the lines of Carroll's Tortoise or, more recently, Gil Harman's work [1986]. But I leave the matter there, turning to the paracomplete case.

[144] Of course, since I'm concentrating entirely on LP- and K_3-based approaches, my claims need to be taken with a grain of salt. It *may* be that some paracomplete approaches can utilize something along the lines of V2*. (I doubt it, but I leave it open.)

[145] In the case of the corresponding V1*, the situation is modeled by a case in which $\nu(A \supset B) = \frac{1}{2}$ (undesignated). Taking an obvious route towards the box, wherein $\nu(\Box A, w) = \mathtt{min}\{\nu(A, w') : Rww'\}$, the same problem is plain.

all, the purported utility of 'strong truth' is precisely to comment on the status of claims for which the dtruth predicate is inadequate. On this route, we have something like

$$\mathsf{Val}(\langle A\rangle, \langle B\rangle) \rightarrow (\mathbb{T}A \supset B)$$

I am not clear whether such an approach will work in standard paracomplete frameworks (e.g., Field), but at least in the (admittedly non-standard) 'paranormal' framework, the approach should work. Whether such an approach will ultimately work, or deliver all that one needs—let alone desires—with respect to 'truth-preservation', is something I here leave open.[146]

7 Leaving Dtruth: Parameters and Contextual Shifts

Until now, I have focused entirely on dtruth and the two main approaches towards resolving dtruth-theoretic paradox. I should note that many (if not most) of the well-known approaches towards truth are not accounts of dtruth: they give up the unrestricted intersubstitutivity that is essential to dtruth. One of the main reasons for largely focusing on dtruth is that, as mentioned in §2, it appears to be the toughest case. If one can resolve the dtruth-theoretic paradoxes in an adequate fashion, then, presumably, one can resolve the truth-theoretic paradoxes for other notions of truth, since the latter notions will be less demanding—e.g., with respect to unrestricted intersubstitutivity or so on. Another reason for focusing on dtruth is that, as mentioned in §1, the other approaches—revision theory and what I here call *parametric* approaches—are well-covered, or at least clearly sketched, in a variety of accessible places. (See §11.) None the less, something should be said about the other approaches, especially since they are both popular and philosophically (as well as logically) interesting.

The most popular (semantic) approaches to truth—at least having left dtruth—-divide into two very broad camps: *parametric* and *revision*. The former camp contains those accounts according to which NTP is answered via some parameter (hidden or otherwise) associated with truth. The latter camp contains those accounts that answer NTP by positing a 'hypothetical' or 'revision rule' character to truth itself. Parametric accounts of truth, at least in broad outline, are perhaps the best known among philosophers. Revision theory is less known among philosophers, despite being an approach that may well afford numerous philosophical applications.

[146] The worry, with respect to Field's framework, is that the proposal (above) calls for a 'unified strong truth' device, something that Field does not have. It *may* be that, in the framework of [Field, 2005b], one could define validity as there being a *good* α such that, for all $\beta > \alpha$, the given argument preserves $\mathbb{D}^{\beta}$-truth. But I don't know whether this will work. In the 'paranormal' (similarly 'residual') account (§4.3), we have a unified 'strong truth' device that ought to do the trick.

My aim in the following 'parametric' sections is entirely informal: I very (very) briefly sketch the ideas of various sub-camps within the broad, parametric camp. For space reasons, I skip over any formal details, leaving such detail to work cited in §11. Similarly, with respect to revision theory, my aim is merely to give a brief, informal sketch of the idea. Fortunately, there are plenty of works that sufficiently discuss both the informal and formal features of the canvassed accounts. (See references throughout, and §11.)

Alfred Tarski [1956] ushered in not only a renewed exploration of truth, but also the broad, and ongoing, parametric tradition towards truth. As such, I quickly sketch Tarski (or, at least, Tarski*), and then turn to other parametric approaches.

7.1 Tarski: language-parameter

As is well known, Tarski [1956] proposed a hierarchical account of *truth for formal languages*. A language is *semantically closed* just when, for all sentences A in the language, there is a name $\langle A\rangle$ of A in the language and all T-biconditionals hold: $\mathsf{T}\langle A\rangle \leftrightarrow A$.[147] *Tarski's Theorem* is that no consistent, first-order language, which is 'syntactically resourceful' (in effect, can sufficiently 'describe' its own syntax), can be semantically closed, at least where $\leftrightarrow$ is material equivalence. The reason arises from Gödel's *diagonal lemma*, which applies to such languages: for any open sentence $A(x)$ with x free, there is a (true) biconditional $B \leftrightarrow A(\langle B\rangle)$, and so Liar-like biconditionals $L \leftrightarrow \neg\mathsf{T}\langle L\rangle$.

Let $\mathcal{L}$ be a first-order, sufficiently syntactically resourceful, language. How can we have a truth predicate for $\mathcal{L}$ such that all T-biconditionals hold and $\mathcal{L}$ is consistent (and, hence, non-trivial)? Tarski's well-known answer invoked a 'richer metalanguage'. In short: do not let T into $\mathcal{L}$! In other words: forget about chasing after semantically closed languages (theories)! The result is that Liar-like sentences never arise, and so truth-theoretic paradox (and, in general, semantic-theoretic) paradox is avoided.

Of course, it is one thing to simply 'ban' Liars from a suitably formal language; it is quite another to so 'ban' from a natural language. Tarski's own views on how, if at all, his proposal might apply to natural languages remain unclear.[148] But let us consider Tarski*, a (fictional) character who proposed a Tarskian, hierarchical account of truth for natural languages.

The proposal, in short, runs as follows. We have some (interpreted) semantic-free 'base language' $\mathcal{L}_0$. No truth-theoretic (or, in general, semantic-theoretic) paradoxes arise in $\mathcal{L}_0$, since the language is devoid of semantic predicates. Let A, B, and C be sentences of $\mathcal{L}_0$, and suppose that A and B are true but C false. Of

[147] This is more accurately put in terms of (the technical, logical sense of) *theories*, but I aim here only to give a sketch. The details are available in many places. See §11 for a few sources.

[148] Tarski seemed to think that truth itself is 'inconsistent' in some fashion, but it is not exactly clear on what he intended. (I have always thought of Tarski as a budding paraconsistentist, but the relevant textual support is weak, to say the least.) For recent discussion of Tarski's views, concerned (unlike here) with precise exegesis, see [Patterson, 2005] and [Soames, 1999] and references therein.

course, since $\mathcal{L}_0$ is devoid of a truth predicate, we can't *explicitly assert in* $\mathcal{L}_0$ that such sentences are true or false. This is the role of the 'next language up', a richer language $\mathcal{L}_1$ that contains a truth predicate T_0^1 for $\mathcal{L}_0$. (Here, I use subscripts and superscripts, respectively, to indicate the languages *for* and *in* which the predicate is introduced.) If the sentences of $\mathcal{L}_0$ are (properly) included among those of $\mathcal{L}_1$, one can consistently have it *in* $\mathcal{L}_1$ that $\mathsf{T}_0^1\langle A\rangle \leftrightarrow A$. (Otherwise, one uses a 'translation' τ, where $\tau(A)$ is the translation of A into an $\mathcal{L}_1$ sentence, and so in $\mathcal{L}_1$ gets the same effect: $\mathsf{T}_0^1\langle A\rangle \leftrightarrow \tau(A)$.) And this gives us what we wanted: we can explicitly state in $\mathcal{L}_1$ the status of our $\mathcal{L}_0$ sentences, for example, $\mathsf{T}_0^1\langle A\rangle$, $\mathsf{T}_0^1\langle B\rangle$, $\mathsf{T}_0^1\langle\neg C\rangle$, and more.

But now what about the status of such $\mathcal{L}_1$ pronouncements? We cannot explicitly assert in $\mathcal{L}_1$ that such pronouncements are true; $\mathcal{L}_1$ contains no truth predicate for itself. This, in turn, is the role of $\mathcal{L}_1$'s metalanguage $\mathcal{L}_2$, wherein we introduce a predicate T_1^2. And so on. In general, a truth predicate T_{n-1}^n of level n applies only to the lower-level sentences, sentences of $\mathcal{L}_m$ for $m \leq n-1$.

Tarski*'s proposal, then, is one that answers NTP by positing a parameter associated with truth: we do not have truth *simpliciter*, but rather only *truth in language such and so*. Tarski*'s parameter, like that of (the real) Tarski, is a language. Non-triviality—indeed, consistency—is achieved by regimenting out Liars.

What about Tarski*'s answer to ECP? In short: Tarski* gives no answer because he rejects the project. Since, according to Tarski*, we do not have semantically significant predicates *for* $\mathcal{L}$ *in* $\mathcal{L}$, we accordingly do not have 'exhaustive characterization' *in* $\mathcal{L}$, for any relevant (sufficiently resourceful) language $\mathcal{L}$. On the other hand, Tarski* certainly does achieve one sense of 'exhaustive characterization', namely, that any suitable *metalanguage* for $\mathcal{L}$ can exhaustively characterize $\mathcal{L}$, and do so consistently. Still, it is important to note that Tarski* rejects ECP, at least in its target sense.

There are various problems associated with Tarski*'s proposal, many of which are well-known. (See §11.) Perhaps the biggest problem is the expressive difficulties associated with the proposal. One wants to be able to (competently) use a truth predicate for just the sorts of broad generalizations for which dtruth was introduced, but in the Tarski* case, things quickly become difficult. For example, in order to (competently, truly) assert that all of So-and-so's claims are true$_n$, one would need to first know that all of So-and-so's claims never reached level n. One could, of course, guess and pick some reasonably high n, but this itself is very difficult, especially if So-and-so is wont to make broad theoretical generalizations concerning language or the like.

A related (and much-discussed) difficultly, made explicit by Kripke [1975], is that some perfectly meaningful—or, at least, seemingly perfectly meaningful—claims wind up being meaningless on Tarski*'s account. Example: suppose that Nixon says (only) *that everything Dean says is false*, and similarly that Dean says (only) *that everything Nixon says is true*. Finding suitable levels according to which such claims are meaningful is difficult on a (consistent) Tarskian* approach, to say the least.

There are other problems (both technical and philosophical) with the Tarskian* account, but enough has been said to move on to another, and very popular, approach—contextualism.

7.2 Contextual truth

So-called contextualist approaches are increasingly dominant today, at least among philosophically inclined logicians and philosophers of language who are content to depart from dtruth.[149] Part of the attraction may be a tied to currently popular contextualist approaches to vagueness. Whatever the reason, the broad contextualist approach to truth is both popular and, at least in basics, fairly straightforward. Because the contextualist approaches are very (very) broad, space considerations allow only a (very) brief, informal sketch of a few leading ideas. The cited works should be consulted both for further, informal remarks and, in particular, the formal account(s).

In short, contextualist approaches maintain that an interpretation of 'true' depends on context, and does so in a way beyond ordinary ambiguity. (Obviously, if A is ambiguous—requiring context to disambiguate—-then an ascription of truth to A will likewise be ambiguous. The context-dependence involved in ordinary ambiguity is not at issue.) In particular, Liars seem to generate inconsistency only when we ignore the implicit contextual parameters associated with truth-ascriptions.

Exactly what sort of context-dependence is relevant to paradox is what distinguishes contextualist theories. In its two most popular forms, such context-dependence is cashed out either with an explicit parameter associated with truth or an implicit parameter essentially involved in (at least Liar-like) truth-ascriptions. I will briefly—and very broadly—sketch these two versions.[150]

Regardless of the version, contextualists—one and all—place a lot of weight on the idea of 'reflecting on paradoxicality' and, in turn, 'moving beyond' such paradoxicality. Consider, for example, the following sentences.

S1. S1 is not true.

S2. S1 is not true.

Contextualists find the following reasoning plausible: S1 is true iff not. Hence (given LEM!), S1 is not true, which is what S2 says, in which case S2 is true.[151]

[149] Actually, there is no reason a contextualist could not acknowledge a dtruth predicate and yet also proceed to acknowledge other (non-transparent) notions of truth—ones not definable in terms of dtruth (and other logical devices). I do not know whether anyone holds such a position, but I think it to be unlikely, since most contextualists are fairly wedded to classical logic. The 'aletheic pluralism' of Crispin Wright [1987; 1992], similarly Michael Lynch [1998], or even that of Michael Dummett [1978], might be a setting for such a position.

[150] For space reasons, I am omitting discussion of other proposals commonly counted as contextualist—e.g., Gaifman, Skyrms, and Koons. See §11.

[151] Strictly speaking, LEM itself can be rejected, but the background logic needs to be 'sufficiently classical'. Contextualists are generally wedded to classical or very-close-to-classical logic.

Notice that the situation is related to 'characterization' and gaps. Consider, for example, the doubly starred sentence:

⋆⋆ The doubly starred sentence is not true.

Assuming the T-biconditionals, which is a common desideratum among contextualists, the doubly starred sentence is true iff not. Hence, given LEM (also a common desideratum), the doubly starred sentence is not true.

But how can such reasoning be sound? In effect, that is the question—or, at least, a chief question—towards which contextualist proposals are directed, and the question on which I'll briefly focus in §7.2–§7.2.

The foregoing examples are closely related to what Bas van Fraassen [1968; 1970] called *the strengthened Liar* phenomenon. For those—call them *caricature gappists*—who wish to both reject LEM *and* assert that some sentences are 'neither true nor false' (i.e., that there are gaps), sentences such as the doubly starred sentence serve as strengthened Liars. The caricature gappist proposes that Liars are gappy, and hence not true. But *that*, by all initial appearances, is precisely what the doubly starred sentence says—namely, that it is not true. And so it is difficult to see how the proposal achieves both consistency and significant characterization (in the sense at issue throughout).[152]

The 'strengthened' phenomenon is not peculiar to caricature gappists; it is quite general, and contextualism, which is—in all standard treatments—intimately tied to classical logic, is often (if not always) motivated by it. For example, suppose that we introduce a semantically significant (unified) predicate Z for purposes of classifying Liars, and that we impose an *exclusive* constraint on our semantically significant categories (i.e., that none of the categories overlap). The proposal, then, is that Liars are (in the extension of) Z. A 'strengthened' Liar immediately emerges:

SL. SL is either Z or not true.

Given the appropriate T-sentence and the assumptions above, one soon arrives at the claim that SL is not true.

The question, as above, is how to make sense of truly—and consistently—asserting that SL is not true, or that S1 is not true, or that the doubly starred sentence is not true, or so on. How can any of this make sense in a (more or less) classical framework in which all T-biconditionals hold? *That*, in effect, is the question.

I argue in my pro-hierarchies paper that there is nothing reall to Simmons' non-hierarchical account that is really not hierarchical. It is in a footnote, but the paper is "Truth, Reflection, and Hierarchies", the goal is to show that hierarchies are non-threatening, natural, and everyone should want them, including Keith.

[152]Of course, at least with *dtruth*, caricature gappism (as it were) makes no sense unless, perhaps as in §5.2, the view is dialetheic (and LEM fails only for *one* negation or etc.).

Indexical truth

One contextualist tradition stands closely with Tarski*. Tyler Burge [1979] and, in much the same spirit, Keith Simmons [1993] maintain that truth is essentially *indexical*.[153] Truth (or, at least, 'is true'), on this story, behaves indexically: the extension of 'true' is determined by context, so that, in effect, we do not have truth simpliciter, but rather *true in context* c.[154]

The idea, in short, is that truth (or 'true') comes equipped with a place-holder or index c for context. In effect, the index is a level in a Tarskian hierarchy, where the given level is determined by context. One context might determine level n, in which case the n-level T-biconditionals are to be used, and in another, level m, whereat the m-level T-biconditionals are in play.[155]

How does any of this answer the questions concerning, for example, S1 and S2? The answer, of course, invokes a change in context, and hence a change in the appropriate indices. In particular, the suggestion is that the (pragmatic) *implicatures* associated in an assertion of S1 are different from those associated with S2. When S1 is asserted in (say) context c, pragmatic principles require that it be evaluated via the c-level T-biconditional:

$$\text{S1 is true}_c \text{ iff S1 is not true}_c$$

But this quickly leads to inconsistency, and hence (more pragmatics) S1 is said to lack c-level truth conditions, in which case (note well) S1 is not true$_c$. But since S1 lacks c-level truth conditions, there are no grounds on which to infer, in turn, that S1 is true$_c$. Recording as much requires 'reflection' in a different context $k \neq c$, wherein S2, which invokes the k-level T-biconditional, may be correctly asserted.

The indexical version of contextualism (be it either Burge's or Simmons') answers NTP by positing a hidden indexical parameter in truth. Classical logic needn't be rejected; one need only recognize that truth is parametric. Non-triviality—indeed, consistency—is achieved in virtue of the contextually relative extensions of 'true'.

[153] Actually, Simmons' theory purports to be entirely non-hierarchical, unlike Burge's. In that respect, there is a significant difference. For space reasons, I will simply sketch the Burgean idea, but one should note that any reference to 'hierarchies' will be rejected on Simmons' account. Another related approach is in [Koons, 1992]. (On the other hand, Michael Glanzberg [2004c] argues that Simmons' apparent claim to achieve a non-hierarchical theory is both unfounded and unmotivated; he argues that hierarchies are non-threatening, and that Simmons' position is one that should—and, in fact, does—embrace hierarchies. But this is beyond the current essay.)

[154] Of course, one might try to define some notion of *truth simpliciter* by quantifying over all contexts: A is true simpliciter iff A is true$_c$ for *all* contexts c. But this is generally not countenanced, since (unrestricted) quantification over all contexts makes for inconsistency, at least given classical (or classical enough) logic—e.g., 'this sentence is false in all contexts' or the like.

[155] This sketch is potentially more misleading than its mere brevity might otherwise afford, because Burge's overall proposal is a pair of proposals, one a formal theory, the other a set of (very complex) pragmatic principles (that are used to interpret the formal theory). For space reasons, I can only wave at the broad picture, leaving the two essential components (formal and pragmatic) in the background.

What about ECP? One perennial worry concerns whether such 'indexicalism' achieves exhaustive characterization in the target sense. If the target sense of 'exhaustive characterization' requires a *unified* predicate in terms of which all Liars are properly classified (in the given language), then the Burge–Simmons approach seems to fail. Without going into details, suppose that we have a predicate—or the same effect, for example via quantification—that applies to all Liars. (Intuitively, one might think along the lines of *not true in any context.*) Let Z be the predicate, and consider SL from §7.2. There seems to be no consistent way of treating such a sentence in the indexicalist's framework. At the very least, some explanation must be given as to why we cannot have such a unified characterization, an explanation that goes beyond pointing to the otherwise resulting inconsistency, an explanation that makes plain why, *despite its apparent intelligibility*, such a notion is incoherent.

Another problem is similar to Tarski*'s (pragmatic) problem. We want to use 'true' for purposes of generalization, but stratification of contexts (on which truth depends) makes this cumbersome, as the pragmatic framework in [Burge, 1979], and similarly (but less explicitly so) in [Simmons, 1993], is quite involved. Indeed, competent, effortless generalizations involving 'true' ultimately require an enormous amount of 'pragmatic know-how', at least on Burge's account. But I will (for space reasons) leave the matter there.[156]

Contextually sensitive quantifiers

The work of Charles Parsons [1974a; 1974b], who is generally credited with initiating the contemporary contextualist tradition, and Michael Glanzberg [2004a; 2004b], who has clarified and extended Parsons' proposal, represents the other popular contextualist tradition. For convenience, I will call the proposal *quantifier-variability.*

The idea, as before, is that truth ascriptions are essentially context-dependent; however, *there is no indexicality involved in truth (or 'is true').* On the quantifier-variability proposal, the context-dependence—the variability—involved in truth ascriptions turns on the essential context-dependence of natural language quantifiers. This is an attractive feature of the proposal. Like the indexicalist account, the background logic is assumed to be entirely classical.[157] But importantly unlike the indexicalist, 'true' is entirely univocal, carrying no hidden parameters at all (or ambiguity, or etc.), at least as applied to 'propositions', which, on the current picture, are the 'chief bearers of truth'.

[156] I should note one thing. Burge (similarly, Simmons) takes his account to be *descriptively accurate* of competent usage of 'true' in English. But at least on Burge's account, such usage, as waved at above, involves an enormous amount of pragmatics—principles that determine the interpretation of (the level of) 'true'. It is implausible, at least on the surface, that such principles are common knowledge, especially since their relevance arises only with respect to paradoxical claims.

[157] Note that, while I am skipping the formal details here, Glanzberg's work focuses heavily on infinitary languages. While this is essential to some of the particular virtues that Glanzberg claims for his account, it can none the less be set aside for purposes of a broad, basic sketch.

Taking propositions as basic—and as the 'things' that are *expressed* by sentences—the relevant T-principles have the following form:

$$\text{If } \langle A \rangle \text{ expresses } p \text{ then } p \text{ is true iff } A$$

And now consider, for example, the doubly starred sentence in §7.2. The verdict, of course, is that the doubly starred sentence does not express a proposition.

If one sees the basic direction of the proposal (thus far), one will immediately foresee a 'strengthened' problem:

$\surd\surd$ The doubly ticked sentence does not express a true proposition.

Given classical logic and the relevant T-principle, the doubly ticked sentence expresses a proposition only if it is both true and not, and so

3. The doubly ticked sentence does not express a proposition.

But, then, familiar (classical) reasoning shows

4. The doubly ticked sentence does not express a true proposition.

And the same familiar (classical) reasoning shows

5. The doubly ticked sentence expresses a true proposition.

What is going on? According to the quantifier-variability proposal, the apparent contradiction is merely apparent, arising from ignoring a contextually shifted quantifier-domain. In particular, there is an implicit existential quantifier at work in the doubly ticked sentence, a quantifier ranging over (a domain of) propositions. The proposal is that both (3) and (5) are true, each being so in virtue of different domains. In the case of (3), the (implicit, existential) quantifier ranges over a domain in which there is no proposition for the doubly ticked sentence to express. In the case of (5), the relevant domain is wider than that of (3), enjoying the requisite proposition.

If one is willing to countenance 'propositions', the quantifier-variability proposal is attractive (at least for those willing to give up dtruth).[158] The story, when filled out (e.g., by Glanzberg), is not only that quantifiers over propositions are context-dependent, but that the background domain of 'truth conditions' (out of which, e.g., 'contents' are often constructed) is similarly subject to contextual expansion—such domains get 'bigger and bigger' without end. Accordingly, the 'shifting' at work in the paradoxes is arguably (and so argued by Glanzberg to

[158] Actually, for those not willing to countenance propositions, Parsons [1983] provides an alternative to the 'proposition' framework, using a truth predicate applying to sentences. (The given predicate may be thought of as being proof-theoretically relativized to a domain of individuals.)

be) a sort common in linguistic practice, in general, and so not some ad hoc posit invoked merely for paradox.[159]

The details, both formal and philosophical, are beyond the current discussion. But the main idea is clear enough. With respect to NTP, the quantifier-variability proposal has it that consistency is achieved in virtue of contextual shifts—*not* shifts of an implicitly parametric truth predicate, but rather quantifier-domain shifts. Liar propositions are certainly in many domains of quantification; it's just that they're not in the domains that would otherwise result in inconsistency.

What about ECP? In short (and very roughly, given the absence of detail), the quantifier-variability proposal rejects ECP, or at least rejects that there is a single context in which we can exhaustively characterize all sentences. The situation is similar in many respects to the indexical version of contextualism (§7.2); the difference is now that there is no 'absolutely unrestricted quantification' with which one would otherwise be able to 'exhaustively characterize' all Liars in the target sense. On the other hand, the quantifier-variability proposal does achieve a great deal. In particular, the proposal seems to be such that for any Liar proposition, there is a context in which the proposition may be truly classified (e.g., as being not true). The price, as mentioned, is the absence of a single context in which to classify all Liars. How steep a price this may be is something that I leave open.[160]

7.3 Situational truth

Jon Barwise and John Etchemendy [1987] propose an account that, in many respects, is similar to the quantifier-variability proposal. Like the latter, truth itself is not indexical in any fashion, and propositions are the chief bearers of truth. Moreover, a sentence (type) may express different propositions depending on context. The difference is that the role of *domains* in the quantifier-variability proposal is now played by *situations*, bits of the world that a statement is about. Hence, propositions are themselves parametric, or at least relative to a given situation. Like the quantifier-variaiblity proposal, 'strengthened' cases—for example, S1 and S2 or the like (see §7.2)—are resolved by recognizing a contextual shift, one that determines a change in situation and, hence, a corresponding change in the propositions expressed in the given contexts.

The Barwise–Etchemendy proposal—which, for convenience, I will call *the situational proposal*—is rich in many ways, both philosophically and logically. A clear

[159]In Glanzberg's work [2004a; 2004b], the sort of contextual shifts witnessed in Liar reasoning are argued to be common contextual shifts—what is 'topical' or 'salient' at a given stage in a discourse—and not merely peculiar to truth-theoretic paradox. It is precisely this broader phenomenon of 'shifting' that, according to Glanzberg, induces the mentioned 'bigger and bigger' expansions of the background domain of 'truth conditions'. (As mentioned above, Glanzberg works with infinitary languages and, by doing so, purports to make explicit the content of various Liar phenomena. In the formal framework, Glanzberg develops levels of a hierarchy that are as semantically closed as Kripke's minimal fixed point.)

[160]Another advantage of the quantifier-variability proposal is that it affords a natural—and entirely analogous—resolution of 'set'-theoretic paradoxes. See [Parsons, 1974b] for discussion. (And see §9 for the reasons behind scare-quoting 'set'.)

account of the proposal is difficult to give without discussing the formal framework, which, in this case, relies on Peter Aczel's ZFC/AFA non-well-founded set theory [1988]. Such a set theory rejects the usual axiom of foundation, and hence allows for 'self-membership' (among other sorts of circularity or, more generally, non-well-foundedness); this allows for (modeling) self-referential propositions in a natural way. For space reasons, I cannot lay out the background set theory or its key theorems, and so cannot go into the formal details of the situational proposal. Instead, I will give a very brief snapshot of the basic idea, leaving cited sources for a fuller and more accurate picture.

As above, the heart of the situational proposal is a contextually sensitive feature of propositions. Two factors combine to determine what proposition is expressed by a sentence on a given use: namely, *demonstrative conventions* and *descriptive conventions*. The former pick out a situation s, and the latter a *type* T of situation. Once these elements are determined, a proposition is at hand, namely, $\{s;T\}$, which is the proposition that s is of type T. This account of propositions and their 'making' is what Barwise–Etchemendy call the *Austinian* account,[161] where the corresponding (Austinian) account of truth is then natural:

$$\{s;T\} \text{ is true iff } s \text{ is of type } T$$

Situations, as above, are bits of the world that a statement is about. To be (slightly) more precise, the *world* consists of *states of affairs*, which, in turn, are built from individuals and properties (or relations); and a set of such states of affairs is a situation. So, in general, states of affairs have the form $\langle R^n, o_1, \ldots, o_n; i\rangle$, where R^n is an n-ary relation (a property if $n = 1$) and i is a polarity, intuitively, a truth value out of $\{1, 0\}$. If $\mathcal{SA}$ is the class of states of affairs, then any situation s is a subset of $\mathcal{SA}$. Moreover, states of affairs yield types of situations: if $\sigma \in \mathcal{SA}$, then $[\sigma]$ is a *type of situation*. Hence, a type T looks like this: $[\langle R^n, o_1, \ldots, o_n; i\rangle]$. Let σ be the state of affairs that determines $T = [\sigma]$. Then the account of truth, above, is equivalent to:

$$\{s;T\} \text{ is true iff } \sigma \in s$$

We say that s is of (simple) type $[\sigma]$ exactly if $\sigma \in s$, and so $\{s;T\}$ is true iff s is of type T.[162]

One might wonder about the role of *the world* in all of this. The role of the world is to distinguish between *accessible* and *inaccessible* propositions. If $p = \{s;T\}$ for some T, then p is *about* s. A proposition is accessible if it is about an *actual* situation, and inaccessible if about a *non-actual* situation. This distinction plays no part in determining the truth (or falsity) of propositions, since there are true (similarly, false) propositions about both actual and non-actual situations. (This is all made abundantly clear in the formal model, but for present purposes, an

[161] This is in contrast to the *Russellian* account, which can also be modeled in the proposed framework. For space reasons, I will restrict the sketch to Austinian propositions.

[162] Barwise–Etchemendy also define molecular types, for example, σ is a disjunctive type $[\bigwedge X]$ exactly if σ is of some type $T \in X$. Similarly for conjunctive types (dual).

informal example will suffice.) Simplifying, if (F, y) is a state of affairs that is not part of the actual world, any accessible proposition that claims (as it were) that y is F, is a false (but accessible) proposition, since any such accessible proposition is one about situations that are not of the required type. On the other hand, if (F, y) is a state of affairs that is part of the actual world, then any accessible proposition that claims (as it were) that y is not F, is likewise false, since the proposition is about an actual situation that is of the wrong type.

The import for simple Liars, for example, *that this proposition is false*, is that they are simply false if about actual situations. For example, a simple Liar proposition $f_s = \{s; [true, f_s; 0]\}$, if accessible, is false because about a situation for which there are no states of affairs of the right type. On the other hand, given situations that contain (e.g.) $\langle true, f_s; 1\rangle$, some Liars are true; it's just that no such situation is (or can be) actual, on pain of triviality.

As with the quantifier-variability proposal, contextual shifts are invoked to explain the sort of reasoning involved in typical 'strengthened' cases—at least, for those, like S1 and S2, that can express propositions. Let f_s, as above, be a Liar about actual situation s. Then f_s is false (for reasons above). But how, then, do we so characterize f_s? We 'expand' s to s', where s' contains the fact (state of affairs) of f_s's falsity. The proposition that we use to characterize f_s is $p = \{s'; [true, f_s; 0]\}$, which is that s_2 is of f_s's type (where f_s is about s), and hence is true because s_2 contains $\langle true, f_s; 0\rangle$. But now s' itself has Liars, for example, $f_{s'} = \{s'; [true, f_{s'}; 0]\}$. But $f_{s'}$ is handled in a similar manner: expand to s'', a situation, unlike s', that contains the fact that $f_{s'}$ is false. Just like the quantifier-variability proposal, the picture is one in which a sequence of propositions emerges, a sequence featuring alternating truth values that reflect contextual shifts: f_s, p, $f_{s'}$, p', and so on. We achieve consistency by switching to different (broader) situations in which the 'previous' statements are evaluated.

The situational proposal, then, is one according to which NTP is answered by relativizing propositions to situations. When one pays heed to the situation a Liar is about, the apparent inconsistency generated by Liars is seen to be merely apparent: the given situations are not of the right type. And the proposal shares many (if not all) of the virtues enjoyed by the quantifier-variability proposal, similarly adverting to a phenomenon—contextual shift—that is arguably a very general one, rather than something peculiar to paradoxes.

With respect to ECP, the situational proposal faces the same problem (or, at least, same sort of problem) confronted by other contextualist proposals. Indeed, though the details are different, the general issue confronting the quantifier-variability proposal emerges for the situational account: in particular, there is no 'global' situation in which we can truly characterize all Liars. Put another way, the world itself is not a situation about which we can consistently talk. (If we could, then there would be global-Liars, and we would be unable to 'expand' to a bigger situation.) But I leave the matter there.[163]

[163] I should note that Barwise–Etchemendy address the given worry in a postscript [1987]. They propose a way in which we can talk about the whole world, that is, have propositions about the

8 Revision Theory

Like the parametric and contextual approaches, revision theorists leave dtruth behind by giving up on the full intersubstitutivity of $\mathsf{T}\langle A\rangle$ and A. But revision theory is significantly different from the foregoing approaches, at least in its philosophical motivation. To fully appreciate the theory (or family of theories), the technical details are important. For space reasons, I cannot go into the technical machinery; instead, my aim is simply to convey the basic idea.

The two pioneers of revision theory are Hans Herzberger and Anil Gupta, who independently came up with the technical ideas in the early eighties. Herzberger's theory [1982] was motivated by an attempt to 'classicalize' Kripke's inductive approach to truth, the aim being a modification affording classical, two-valued valuations, one according to which Liars (or the like) were neither gaps nor gluts, but rather *unstable*. The given 'instability' is reflected in the manner in which sentences are interpreted. Specifically, sentences are evaluated classically but via a sequence of stages; on some such stages, paradoxical sentences are true, and on others false. One of the chief aims of Herzberger's theory is to characterize just such patterns of instability.

Gupta [1982] and, in turn, Gupta–Belnap [1993], motivate a revision theoretic approach to *truth* as an instance of a much broader revision theoretic approach to *definitions*, in general. For present purposes, I will focus on the Gupta–Belnap proposal [1993], which provides a much broader philosophical picture than found in Herzberger.[164]

On the traditional account of definitions, a proper definition is neither *circular* nor *creative*. A definition is circular if the definiendum appears in the definiens. A definition is creative if it yields proofs of claims that are not reducible to definiendum-free claims. While revision theory maintains the stricture against creativity, the proposal breaks from tradition by allowing circular definitions. Indeed, in large part, the point of revision theory, at least qua theory of definitions, is to allow for meaningful, useful circular definitions. Gupta–Belnap propose to preserve classical logic, and so we assume as much in the background.[165]

Consider an example that nicely illustrates the general idea [Gupta and Belnap, 1993]. Suppose that we are given the following definition of G, and that our domain

world at large. The result is that we can sometimes talk about the whole world, but any Liar-sentence expresses a proposition only about a proper part of the world. On the surface, this is philosophically suspect, since once we are able to talk about the whole world, it is very difficult to see how global-Liars do not emerge. Indeed, once talk of the whole world is available, one could introduce a convention that indicates global talk, some symbol that indicates a global reading of the proposition.

[164] This is not a criticism of Herzberger's work. The point is that, for present purposes, the broader philosophical perspective is appropriate.

[165] It should be noted that revision theory, at least the basic idea, is compatible with many non-classical logics. Indeed, see the 'revision flavor' of Field's conditional, which is clear in the 'restricted semantics' (§4.2).

is $\{a, b, c, d\}$.

$$Gx =_{df} (Fx \wedge Hx) \vee (Fx \wedge \neg Hx \wedge Gx) \vee (\neg Fx \wedge Hx \wedge \neg Gx)$$

Suppose, too, that we have the following information:

i. a and b are (in the extension of) F.

ii. c and d are not F.

iii. a and c are H.

iv. b and d are not H.

How might we determine the Gs? How, in other words, might we go about discovering the extension of G? The obvious problem confronting our definition of G is that it is circular. Figuring out the extension of G requires, it seems, already knowing the extension of G. So it appears.

Revision theory enters with a simple (but powerful) solution: posit a *hypothetical* extension and proceed to calculate the extension of G on the basis of the hypothesis! Let h_0 be your initial such hypothesis. Your subsequent calculation will yield an extension of G *relative to* h_0, an extension that, depending on h_0, may well differ from h_0, and hence *revise* your initial hypothesis.

Consider, for example, an initial hypothesis h_0 according to which $\mathcal{G}$, the extension of G, is $\emptyset$. With respect to object a, we immediately have it that $a \in \mathcal{G}^{h_0}$, that is, that a is in *the extension of G relative to* h_0, since, given (i) and (iii), we have the truth of $Fa \wedge Ha$, and so the a-instance of the entire (disjunctive) definiens, which is evaluated classically, is true. Likewise, it is clear that $c \in \mathcal{G}^{h_0}$, since, given (ii) and (iii), we have the truth of $\neg Fc \wedge Hc \wedge \neg Gc$, and so the truth of the c-instance of the entire definiens. Similar calculation shows that neither b nor d is in $\mathcal{G}^{h_0}$. Hence, relative to our initial (in this case, null) hypothesis h_0, we have come upon an extension for G, one that is relative to, and distinct from, h_0, namely $\{a, d\}$.

But what now? After all, h_0 was just a blind guess! The next step, according to revision theory, is to start again, but this time running with the 'discovered' extension $\mathcal{G}^{h_0}$ as one's hypothesis (i.e., hypothetical extension). So, in the 'second run', h_1 is the hypothesis that $\mathcal{G} = \{a, d\}$. Simple calculation based on h_1 and (i)–(iv) will yield an extension $\mathcal{G}^{h_1}$, the extension of G *relative to hypothesis* h_1. And so on.

Thinking of n as the *stage* at which h_n is posited, the general recipe—though eventually (but not here) carried into the transfinite—is basically as follows. One begins with an initial hypothesis h_0 to get an extension $\mathcal{E}^{h_0}$ that is relative to (but frequently different from) h_0. At stage 1, one's hypothesis h_1 is that the relevant extension is $\mathcal{E}^{h_0}$. At stage 2, one's hypothesis h_2 is that the relevant extension is

$\mathcal{E}^{h_1}$. In general, one's hypothesis at stage $n+1$ is that the relevant extension is $\mathcal{E}^{h_n}$, an extension 'discovered' by calculation relative to h_n.[166]

The foregoing recipe, the process, generates a so-called *revision sequence*, which is a sequence of 'candidate extensions' for the relevant predicate. Each such revision sequence is *relative to the initial, stage-one hypothesis.* Consider, again, the case of G above. For $n > 0$, let $\langle n, x\rangle$ represent *stage* n at which the extension of G relative to h_{n-1} is x. (In other words, x is the extension one gets at stage $n+1$ by using the n-stage extension as initial input.) Then the revision sequence, *relative to our initial (null) hypothesis* h_0, looks as follows.

$$\langle 0, \emptyset\rangle, \langle 1, \{a, c\}\rangle, \langle 2, \{a\}\rangle, \langle 3, \{a, c\}\rangle, \langle 4, \{a\}\rangle, \ldots$$

This sequence, as said, is relative to our given h_0, in this case, an initial hypothesis according to which nothing is G. But, of course, with a domain of four objects, there are sixteen possible starting points, sixteen possible initial hypotheses (only one of which is null). Canvassing *all* such resulting sequences exhibits the sort of 'patterns of (in-) stability' that revision theorists seek to illuminate. Two notions are central:

S. Sentence B is *stable* if, after various stages, B forevermore enjoys the same value (either true or false).

C. Sentence B exhibits *convergence* if B stabilizes on the same value in all sequences.

In the current case, given the sixteen possible starting points and the resulting revision sequences based on (i)–(iv), the following patterns emerge:[167]

R1. Object a. If h_0, the initial (not necessarily null) hypothesis has it that $a \in \mathcal{G}$, then $a \in \mathcal{G}^{h_n}$ for all n. If h_0 has it that $a \notin \mathcal{G}$, then $a \in \mathcal{G}^n$ for all $n > 0$. Accordingly, the sentence Ga is eventually forevermore true regardless of starting point, and hence both *stable* and *convergent.*

R2. Object b. If h_0 is such that $b \in \mathcal{G}$, then $b \in \mathcal{G}^{h_n}$ for all n, in which case Gb stabilizes on all such sequences (sharing the 'positive' b-hypothesis). Similarly, if h_0 has it that $b \notin \mathcal{G}$, then $b \notin \mathcal{G}^{h_n}$ for all n, in which case, again, Gb stablizes. So, Gb is stable on all initial hypotheses, but, since it doesn't stabilize on the same value in all sequences, it fails to be convergent.

R3. Object c. There is no stability with respect to Gc, and hence no convergence.

R4. Object d. This is exactly like the case of a, except that Gd is converges on *falsity*, rather than *truth.*

[166] Because I am omitting the formal details, I am also omitting discussion of the revision process defined over transfinite ordinals. What should be noted is that different revision theories largely differ with respect to their treatment of limit ordinals and, as a result, how 'categorical' judgments about extensions are garnered. See §11 for relevant works.

[167] Such patterns are best conveyed graphically, but for present purposes I merely sketch the import.

With such patterns at hand, and the corresponding features of stability and convergence, revision theory finally arrives at so-called categorical verdicts, or 'absolute values', as one might say. After all, if the revision process is to be of any use, we want to arrive at a *single* judgment about (say) G's extension, as opposed to a bunch of judgments based entirely on arbitrary starting points. Not surprisingly, this is precisely the role of *convergence*. In the case at hand, we conclude that a is categorically G and that d is categorically non-G. And those, at least in the given case, exhaust the available categorical judgments. Nothing categorical can be said of c, as it remains simply *unstable*. Similarly, nothing categorical can be said of b, since, despite stability in each sequence, there is no convergence. Convergence gives us *a single value irrespective of starting point*. Failing that, it would be unreasonable to make a categorical judgment. And so it is with revision theory.

What does any of this have to do with *truth*? The answer, in short, is that truth itself is said to be a circular concept. In particular, Gupta–Belnap posit an ambiguity in the usual T-biconditionals. Such biconditionals can be read either *definitionally* or *materially* (i.e., material equivalence). Each such biconditional, when *definitionally* read, has the form

$$\mathsf{T}\langle A_i \rangle =_{df} A_i$$

for $i \in \mathbb{N}$. But, then, since A_i, for some i, may contain T, the resulting definition of T, namely,

$$\mathsf{T}(x) =_{df} (x = \langle A_0 \rangle \wedge A_0) \vee (x = \langle A_1 \rangle \wedge A_1) \vee (x = \langle A_1 \rangle \wedge A_1) \vee \ldots$$

will be circular, the definiendum appearing in the definiens.[168]

The basic picture, then, is clear. Our definition of truth—which relies on the definitional T-biconditionals—is circular. Revision theory is a theory of circular definitions. The proposal is that *the essential character of truth*, its essential definition, is a *rule of revision*, a rule that tells us how to go from 'initial hypotheses' to new hypotheses about the extension of 'true', a rule that ultimately yields useful *categorical judgments* about truth—about the extension of 'true'. Just as in the (abstract) example of G, so too with truth: some sentences will converge, others stabilize but not converge, and others never stabilize at all. Paradoxical sentences, at least ones that would otherwise result in inconsistency, fail to stabilize.

The exact judgments concerning Liars and other 'ungrounded' sentences depend on the particular revision theory. (See §11 for relevant works.) But the general proposal, as above, is clear. With respect to NTP, for example, the revision theoretic proposal is that non-triviality—indeed, consistency—is achieved in virtue of truth's 'revision-rule character', one that results in certain sentences (e.g., Liars) being unstable.

Unlike the traditional account of definitions, according to which circularity often results in inconsistency, appropriate revision rules serve to allow meaningful

[168]I should mention, but (for space) cannot discuss, that Gupta–Belnap prove various results about the behavior of $=_{df}$ (as it were) and $\equiv$ (material equivalence), showing that they are the same except in abnormal cases. See [Gupta and Belnap, 1993] for full discussion.

and *useful* circular definitions without inconsistency. A nice feature of revision theoretic approaches is that, as in the quantifier-variability and situational approaches, a *general* phenomenon is invoked to explain paradox, rather than some ad hoc feature peculiar to paradox. In the present case, the general phenomenon is circular definitions.

What about ECP? In short, revision theorists reject the project. Any version of revision theory (as far as I can see) will invoke semantically significant categories that, were they introduced into the (object-) language, would result in inconsistency. Consider, for example, a typical 'strengthened' case like 'this sentence is not true or unstable'. This sentence poses problems. Of course, one could, as suggested in [Gupta and Belnap, 1993], introduce 'unstable' into the language and, in turn, have a richer metalanguage that categorized the problem sentence as (say) 'super-unstable' or the like. But, again, nothing close to the (target sense) of exhaustive characterization is achieved. On the other hand, it may well be, as argued in [Gupta and Belnap, 1993], that for any stage s of (say) English, there is a stage s' in which s-English is exhaustively characterized. What one won't have—at least not consistently—is some stage s such that s-English exhaustively characterizes s'-English *for all* s'. As with other proposals, I leave the philosophical cost-benefit analysis of the matter open.[169]

9 Set-theoretic paradoxes?

For space reasons, this essay has not covered the wide range of so-called semantic paradoxes that, in many ways, are closely related to truth-theoretic paradox—for example, denotation, satisfaction, reference, etc. Instead, the focus has been on the Liar, and a limited number of approaches to the Liar.[170] But something should be said about so-called set-theoretic paradoxes, which, despite Ramsey's division [1925], are often thought to be in kind with the Liar. For present purposes, I will only give a brief comment.

Russell's paradox is often thought to be the set-theoretic counterpart of the Liar. The naïve comprehension scheme, according to which, for any open sentence $A(x)$, there is a set $\mathcal{Y}$ such that $o \in \mathcal{Y}$ iff $A(o)$, for any object o. Russell's paradox invokes $x \notin x$. By the comprehension scheme, we have some set $\mathcal{R}$ comprising all and only the objects that are 'non-self-membered', that is, the objects satisfying

[169] I should make it plain that Gupta–Belnap [1993] argue that ECP, inasmuch as its target sense overlaps with what they call 'semantic self-sufficiency', is unfounded. My own thinking is that we *start* with the strong appearance of 'exhaustive characterization', that (e.g.) natural languages are exhaustively characterizable in terms of unified (semantically significant) devices. If we cannot make sense of how this could be—if, that is, we cannot answer ECP adequately—then we should look for alternatives, perhaps along the lines of the 'stages of English' suggestion above (where the order of quantifiers is critical). But, for repeated space reasons, I must leave the matter there.

[170] Indeed, I have not covered even the wider family of truth-theoretic paradoxes, those paradoxes (e.g., Knower, Necessity, etc.) that involve 'truth-entailing notions'.

$x \notin x$. But, then, $\mathcal{R} \in \mathcal{R}$ iff $\mathcal{R} \notin \mathcal{R}$. The resemblance to Liars and the T-schema is obvious.

What should be noted is that, in effect, *Russell's paradox* amounts to two paradoxes, one for *set theory*, one for *semantics*. By my lights, mathematics is free to simply axiomatize away its version of Russell's paradox, as is standardly done. Sets were introduced within and for mathematics, a role that need not be constrained by our 'intuitions' about natural language. So long as the given role is sufficiently served by 'axiomatically harnessed' sets, then so be it. Of course, mathematicians frequently seek the most natural or 'beautiful' route in their endeavors, and perhaps standard axiomatized approaches towards sets are not as natural as one might like. But that is a separate issue. The main goal is to find entities that play the given mathematical role.

Semantics, on the other hand, seems to be a different matter. What semantics needs is a theory of (what I will call) *semantical properties*.[171] Semantical properties are entities that play a particular semantic role, in particular, those entities corresponding to each meaningful predicate in the (given) language. The role of semantical properties is essentially given by the naïve semantical schema: namely, that for any object o and predicate $Y(x)$, there is a semantical property $\mathcal{Y}$ such that o exemplifies $\mathcal{Y}$ iff $Y(o)$ is true. This principle is fundamental to semantics, at least in practice. But the principle confronts Russell's paradox for *semantics*, in particular, the instance invoking x *does not exemplify* x.

Perhaps one reason that the two paradoxes are not distinguished is that semanticists simply borrowed the mathematical sets to play the role of semantical properties, and the history of (at least contemporary, formal) semantics has gone along with the borrowed entities. But why think that the mathematicians' sets will sufficiently play the role of semantical properties? There is no a priori reason to think as much. Indeed, inasmuch as semantics needs its unrestricted naïve semantical schema (above), there is reason to be pessimistic.

A solution to the Liar, at least by my lights, need not thereby be a solution to *mathematical set*-theoretic paradoxes; those paradoxes may be solved in whatever way suits mathematics. Of course, if mathematics remains classical, then one's proposed 'all-purpose logic' will need to have classical logic as an extension. The point is that Russell's set-theoretic paradox needn't be an issue in one's resolution of the Liar or truth-theoretic paradoxes, in general. On the other hand, one's resolution of the Liar is constrained, to some degree, by Russell's semantics-theoretic paradox—the semantical properties paradox. At the very least, solutions to the Liar that require rejecting the unrestricted semantical schema would seem to carry a prima facie blemish.

As it turns out, accommodating the unrestricted semantical schema usually goes hand in hand with achieving a suitable conditional, one for which $A \rightarrow A$ is valid,

[171] I used to use the term 'semantical extensions', but Hartry Field (in conversation) convinced me that 'extensions' is apt to mislead, perhaps suggesting the target entities are *extensional*. I use Field's suggested term 'properties', which is more natural.

$\to$ detaches, and Curry problems avoided.[172] Accordingly, some of the theories sketched above, notably the paracomplete and paraconsistent, naturally afford a solution to Russell's semantical paradox.[173]

10 Revenge—at last

Charges of *revenge* are ubiquitous in the Liar literature. The loose metaphor aptly conveys the target phenomenon, fuzzy as the phenomenon may be. Dramatically put, Liars attempt to wreak inconsistency (or triviality) in one's language. One way of avoiding the Liar's intentions is to kill the Liars, much along Tarski*'s line or lines (not discussed here) that treat them as meaningless. But killing Liars, at least natural-language Liars, carries costs, as mentioned in §7.1. The more humane approach, exhibited by all of the canvassed approaches (except Tarski*), is to let Liars live but frustrate their aim of generating inconsistency (or triviality). And this is where the Liar's revenge is thought to emerge. In short, the Liar's aims are safely frustrated only at an expressive cost. If the Liar can't have what she wants, she'll enlist 'strengthened' relatives to frustrate your wants, in particular, your expressive wants.

So goes the drama. In less dramatic terms, *revengers* charge that a given (formal) theory of truth fails to be explanatorily adequate with respect to its target phenomena. In all cases, the target phenomena include at least the consistency—or, more broadly, non-triviality—of truth, *truth* in our *real* language.[174] The revenger charges that the formal 'model language' achieves consistency (or non-triviality) only in virtue of expressive inadequacies. In particular, some semantic

[172] I should note that the semantical properties version of Curry is directly analogous to the truth-theoretic version. Let ϵ be the exemplification relation, $[x : A(x)]$ the semantical property corresponding to $A(x)$, and $x\epsilon[y : A(y)] \leftrightarrow A(x)$ the semantical property schema. Curry is generated via the predicate $x \in x \to \bot$, yielding $[x : x \in x \to \bot]$, where $\bot$ is either explosive or plainly false.

[173] For concrete examples, see [Field, 2004] and [Priest, 1987]. I should note that Priest [1987] argues that any adequate solution to the Liar must be a solution to Russell's set-theoretic paradox. I reject the argument, but the matter ultimately turns on whether there is a sustainable distinction between semantical properties and mathematical sets. Priest (in conversation) thinks not, but also thinks that mathematics itself is ultimately inconsistent, thereby calling for a paraconsistent logic. These issues, needless to say, are (far) beyond the scope of this essay.

[174] All of this assumes, as I have throughout, that we are dealing with proposals that purport to be *descriptively adequate* with respect to various target phenomena. Those, like Vann McGee [1991], who think that our real language is inconsistent—due to Liars—and, in turn, proceed to offer a chiefly *prescriptive* (consistent) remedy, are not covered in this discussion. With respect to the current issue, McGee's position is somewhat curious. If, as McGee [1991] concludes (on the descriptive side), English is inconsistent, then why think that its logic is explosive (e.g., classical)? Presumably, it is because the logic is explosive that a prescriptive remedy is required. But a more reasonable hypothesis, at least on the surface, is that the logic is paraconsistent, at least if—as McGee maintains—our 'naïve theory of truth' is inconsistent. But if the logic is paraconsistent, then why do we need a consistent remedy? But I must leave the matter there. (I should mention that McGee's work, which, only for space reasons, I have not been able to cover, is among the pioneering works in the 'determinate truth' tradition.)

notion $\mathcal{X}$, which is expressible in our real language, is, *on pain of inconsistency (or triviality)*, inexpressible in the (formal) model language.

Doing formal semantics is modeling work, especially when one is engaged in philosophical semantics or philosophical logic, where the aim, in general, is to illuminate the structure and logic of philosophically significant notions—semantic notions (truth, language, etc.), metaphysical notions (worlds, parts, etc.), perhaps ethical notions, or so on. With respect to formal theories of truth, the aim is no different; one aims to give a model that sufficiently illuminates various features of truth. Of course, in the case of truth, one aims to give a *model* in more than the ordinary, heuristic sense (e.g., exemplar); one also aims to establish *consistency* or, more broadly, *non-triviality*.

In giving a formal theory of truth, one does not directly give a theory of *truth*; rather, one gives a theory of *truth-in-*$\mathcal{L}_m$, for some formal 'model language' $\mathcal{L}_m$. By endorsing a formal theory of truth, one is endorsing that *truth* itself is relevantly like *that*, like truth-in-$\mathcal{L}_m$, at least with respect to various phenomena in question—for example, truth's consistency (or non-triviality), its logic, etc. Revengers charge that truth-in-$\mathcal{L}_m$ is insufficiently like truth for purposes of explaining the target phenomena.

For simplicity, let $\mathcal{L}_m$ be a given formal model language for $\mathcal{L}$, where $\mathcal{L}$ is our target, real language—the language features of which (e.g., consistent truth) $\mathcal{L}_m$ is intended to explain or otherwise illuminate. Let $M(\mathcal{L}_m)$ be the metalanguage for $\mathcal{L}_m$, and assume, as is typical, that $M(\mathcal{L}_m)$ is a fragment of $\mathcal{L}$. Then various (related) recipes for revenge run roughly as follows.[175]

Rv1. Recipe One.

- Find some semantic notion $\mathcal{X}$ that is used in $M(\mathcal{L}_m)$ to classify various $\mathcal{L}_m$-sentences (usually, paradoxical sentences).
- Show, in $M(\mathcal{L}_m)$, that $\mathcal{X}$ is not expressible in $\mathcal{L}_m$ lest $\mathcal{L}_m$ be inconsistent (or trivial).
- Conclude that $\mathcal{L}_m$ is explanatorily inadequate: it fails to explain how $\mathcal{L}$, with its semantic notion $\mathcal{X}$, enjoys consistency (or, more broadly, non-triviality).

Rv2. Recipe Two.

- Find some semantic notion $\mathcal{X}$ that, irrespective of whether it is explicitly used to classify $\mathcal{L}_m$-sentences, is expressible in $M(\mathcal{L}_m)$.
- Show, in $M(\mathcal{L}_m)$, that $\mathcal{X}$ is not expressible in $\mathcal{L}_m$ lest $\mathcal{L}_m$ be inconsistent (or trivial).

[175]This is not in any way an exhaustive list of recipes! (A related route, for example, is pursued by Glanzberg [2001], who argues that any 'good' theory of truth cannot be *applied* without invoking devices that, were they 'allowed' in the theory, would result in inconsistency or triviality. A proof-theoretic version of this route towards 'revenge' is in [Glanzberg, 2004c].)

- Conclude that $\mathcal{L}_m$ is explanatorily inadequate: it fails to explain how $\mathcal{L}$, with its semantic notion $\mathcal{X}$, enjoys consistency (or, more broadly, non-triviality).

Rv3. Recipe Three.

- Find some semantic notion $\mathcal{X}$ that is (allegedly) in $\mathcal{L}$.
- Argue that $\mathcal{X}$ is not expressible in $\mathcal{L}_m$ lest $\mathcal{L}_m$ be inconsistent (or trivial).
- Conclude that $\mathcal{L}_m$ is explanatorily inadequate: it fails to explain how $\mathcal{L}$, with its semantic notion $\mathcal{X}$, enjoys consistency (or, more broadly, non-triviality).

A *revenger*, as above, is one who charges 'revenge' against a formal theory of truth, usually along lines above. The charge is that the model language fails to achieve its explanatory goals. Suppose that $\mathcal{L}_m$ is some formal language that enjoys its own truth predicate. Suppose, further, that consistency (or, for paraconsistentists, non-triviality) is established for the given language (and the resulting truth theory). A *revenger* charges that $\mathcal{L}_m$ fails to explain its target explanada, that, in particular, its consistency (or non-triviality) is achieved in virtue of expressive inadequacies. In general, the revenger aims to show that there's some sentence in $\mathcal{L}$ (our real language) that *ought* to be expressible in $\mathcal{L}_m$ if $\mathcal{L}_m$ is to achieve explanatory adequacy.

How ought one to reply to revengers? The answer, of course, depends on the details of the given theories and the given charge of revenge. For present purposes, without going into such details, a few general remarks can be made.

First, the weight of (Rv1) or (Rv2) depends on the sort of $\mathcal{X}$ at issue. As in §4.1–§5.1, if $\mathcal{X}$ is a model-relative notion constructed in a *proper fragment of* $\mathcal{L}$, then the charge of inadequacy would seem to carry little weight. In particular, as mentioned in §4.1–§5.1, if classical logic extends that of $\mathcal{L}_m$, then there is a clear sense in which you may 'properly' rely on a classical metalanguage in constructing $\mathcal{L}_m$ and, in particular, truth-in-$\mathcal{L}_m$. After all, you endorse that $\mathcal{L}$, the real, target language, is non-classical but enjoys classical logic as a (proper) extension. Accordingly, while details need to be given, there is nothing suspect about relying on an entirely classical (proper) fragment of $\mathcal{L}$ in which to construct your model language.[176] But in such a context, it is hardly surprising that $\mathcal{X}$, being an entirely classical notion (at least over $\mathcal{L}_m$), would bring about inconsistency or, worse, triviality, if we were to express it in $\mathcal{L}_m$.[177]

Because classical logic is typically an extension of the logic of $\mathcal{L}_m$, the point above is often sufficient to undermine the revenger's claims, at least if the given

[176] Note too that if the given fragment is 'necessarily classical' in some relevant sense, then inferences within the given (proper) fragment of $\mathcal{L}$ may turn out to be valid simpliciter—depending on how this is cashed out. Otherwise, some restricted but, as far as I can see, entirely appropriate notion of validity will be in play for the 'metalanguage inferences'.

[177] Hartry Field [2003b] discusses this point with respect to his particular proposal. For a fuller discussion, see [Field, 2005b].

recipe is (Rv1) and (Rv2). What the revenger must show, ultimately, is not the unsurprising result that a model-relative $\mathcal{X}$ is expressible in $M(\mathcal{L}_m)$ but not in $\mathcal{L}_m$; she must show that some relevant non-model-relative notion is expressible in $\mathcal{L}$ but, on pain of inconsistency (or triviality), inexpressible in $\mathcal{L}_m$. And this task brings us to (Rv3).[178]

Recipe (Rv3) is perhaps what most revengers are following. In this case, the idea is to locate a non-model-relative (semantic) notion in $\mathcal{L}$ and show that $\mathcal{L}_m$ cannot, on pain of inconsistency (or triviality), express such a notion. As may be evident, the dialectic along these lines is delicate. Suppose that Theorist proposes some formal theory of truth, and Revenger, following (Rv3), adverts to some non-model-relative (semantic) $\mathcal{X}$ that (allegedly) is expressible in $\mathcal{L}$. If, as I'm now assuming, Theorist neither explicitly nor implicitly invokes $\mathcal{X}$ for purposes of (semantic) classification, then Revenger has a formidable task in front of her. In particular, Revenger must, without begging the question, show that $\mathcal{X}$ really is an intelligible notion of $\mathcal{L}$.

An example, relevant to the paracomplete and paraconsistent cases (see §4–§5), might be useful. Let us say that a language contains an *absolute exclusion device* $\heartsuit$ exactly if *both* of the following hold for the given language.

ed1. $\Vdash \heartsuit A \vee A$

ed2. $\heartsuit A, A \Vdash B$

Against a paracomplete or paraconsistent proposal, a revenger, following (Rv3), will not make the mistake of pointing to some model-relative exclusion device. Rather, the (Rv3)-type revenger might maintain that $\mathcal{L}$, our real (and target) language, enjoys such an absolute exclusion device. If the revenger is correct, then standard paracomplete and paraconsistent proposals are inadequate, to say the least. But the issue is: why think that the revenger is correct? Needless to say, argument is required. What makes the matter delicate is that many arguments are likely to beg the question at hand. After all, according to paracomplete and paraconsistent theorists, what the Liar teaches us is that there is no absolute exclusion device. Accordingly, the given revenger cannot simply point to normal evidence for such a device and take that to be sufficient. On the other hand, if the proposed theory cannot otherwise explain—or, perhaps, explain away—normal evidence for the (alleged) device, then the revenger may make progress. But the situation, as said, is delicate.

The burden, of course, lies not only on Revenger; it also lies with the given Theorist. Paracomplete and paraconsistent theorists must reject the intelligibility of any absolute exclusion device. But inasmuch as such a notion is independently plausible—or, at least, independently intelligible—such theorists carry the burden of explaining why such a notion appears to be intelligible, despite its ultimate

[178] There is (obviously) a lot more that can be said about (Rv1) or (Rv2), but for space reasons I briefly turn to (Rv3).

unintelligibility. (E.g., we are making a common, reasonable, but ultimately fallacious generalization from 'normal cases' to all cases, or some such mistake.) Alternatively, such theorists might argue that, contrary to initial appearances, the allegedly intelligible notion is in fact rather unclear; once clarified, the problematic features disappear. (E.g., one might argue that the alleged notion is a conflation of various notions, each one of which is intelligible but not one of which behaves in the alleged, problematic way.) Whatever the response, theorists do owe something to (Rv3)-revengers: an explanation as to why the given (and otherwise problematic) notion is unintelligible.[179]

I have hardly scratched the surface of *revenge* in the foregoing remarks. The phenomenon (or, perhaps more accurately, family of phenomena) has in many respects been the fuel behind formal theories of truth, at least in the contemporary period. In many respects, at least in my opinion, a clear understanding of revenge is a pressing and open matter. What, exactly, is revenge? How, if at all, is it a serious problem? Is the problem *logical*? Is the problem *philosophical*? And for what end, exactly, is the alleged problem a problem? How is it related to ECP, NTP, or the likes of 'semantic closure'?

Answers to some of the given questions, I hope, are clear enough in foregoing remarks. But answers—clear answers—to many of the questions remain to be found. Until then, full evaluation of current theories of truth remain out of reach.

11 Further reading

This section is intended as a brief—but in *no* way exhaustive—suggestion towards further reading on the main topics. The bibliographies in the given works contain a broader guide. In the case of edited collections, I try to list fairly recent works, which are likely to be widely available. (Again, the bibliographies in the given works will point to many other works of direct relevance.) Unless otherwise stated, *collections* (as used here) are volumes that contain papers by various authors, rather than a single author's collected papers.

11.1 Truth and Paradox

- [Beall, 2003]. While some of the papers in this collection focus chiefly on soritical paradoxes and issues of vagueness, some of the others focus squarely on semantical or logical paradox, with a few papers tying the two kinds of paradox together.

- [Beall and Armour-Garb, 2005] is a collection that focuses squarely on dtruth and dtruth-theoretic paradox, and in particular on the prospects of achieving 'deflationistically acceptable' solutions to such paradox.

[179] For a recent example of attempting to meet this burden, see [Field, 2005b].

- [Chapuis and Gupta, 2000] is a collection that focuses on revision-theoretic approaches to paradox, and also contains discussion of extending revision theory in other philosophical and logical directions.
- [Martin, 1970] is a collection that represents approaches to truth-theoretic paradox in the 1960s and early 70s.
- [Martin, 1984] is a collection that represents approaches that, from the late 1970s to early 80s, have emerged as philosophically popular and influential in the area.
- [Visser, 2004] is a much broader, detailed discussion of many of the approaches discussed in the current essay, with a focus on both philosophical and, especially, the mathematical aspects of the given theories. (The given bibliography is also quite valuable.)

11.2 Dtruth and 'nature' issues

- [Armour-Garb and Beall, 2005] is a collection of papers representing the various deflationary conceptions of truth. (Note that 'deflationism' is really just a catch-all phrase for a family of closely related theories or, more broadly, views.)
- [Blackburn and Simmons, 1999] is a collection of papers that give a flavor of the 'nature' debates. (The introductory essay pushes the relevance of truth-theoretic paradox to 'nature' questions.)
- [David, 1994] is a monograph that provides an extensive examination of both correspondence theories and deflationary theories of truth.
- [Greenough and Lynch, 2005] is a collection that, in many ways, ties approaches to—the nature, if any, of—truth to broader philosophical and methodological issues.
- [Halbach and Horsten, 2002b] is a collection that discusses a wide variety of truth-related topics; it contains both 'nature' topics and discussion of truth-theoretic (and, in general, semantic) paradox.
- [Künne, 2003] is a monograph that examines a variety of popular or historically significant positions on 'nature' issues. (Though it is not a history, the historical information in this book is valuable.)
- [Lynch, 2001] contains essays on the 'nature' debates, including representatives of *correspondence theory*, *coherence theory*, *pragmatism*, *aletheic pluralism*, and various *deflationary* positions.
- [Schantz, 2002] is a collection covering a wide range of 'nature' issues, as well as containing a few essays examining Tarski's so-called Convention T.

11.3 Paraconsistency

- [Priest, 2002] is a wide-ranging survey of paraconsistent logics and their applications.
- [Priest *et al.*, 2004] is a collection that focuses on dialetheic approaches to paradox and paraconsistency, more generally.
- [Norman *et al.*, 1989] is a collection that focuses on paraconsistency, quite generally—touching on various logics and philosophical applications of such (paraconsistent) logics.
- [Priest, 2001] is a monograph that discusses the motivations of dialetheism found in various branches of philosophy (and mathematics); its discussion of semantical paradox is directly relevant to Priest's version of dialetheism.

11.4 Paracomplete

As mentioned in the main discussion, paracomplete approaches are the dominant trend, at least with respect to dtruth, but also popular for other (less transparent) notions of truth. As a result, many—if not all—of the major works cited in the main discussion provide discussion and further references.

11.5 Tarski

Tarski's own work [1944; 1956] is quite readable (and worth reading). Discussion of Tarski's (or, at least, Tarski*'s) approach to truth-theoretic paradox may be found in almost all of the works cited in §11, but see especially the two Martin collections (see §11.1) and references therein.

11.6 Contextual

In addition to the positions sketched in the main discussion, a few other 'contextualist' positions are prominent. (And for related works, see the respective bibliographies.)

- [Koons, 1992] provides a combination of 'situational truth' and Burgean 'indexical truth', with the main advantage being that only two levels are required for 'indexing', as opposed to infinitely many (as on the full Burgean proposal).
- [Gaifman, 1992] offers an approach often classified as a contextualist approach to truth. Semantic evaluation generates a 'network', and the interpretation of an utterance of 'true' depends on the place of the utterance in the given network.

- [Skyrms, 1970; Skyrms, 1984] provides an 'intensional' approach to truth, wherein substitutivity of identities fails. (The overall approach is often seen as 'contextual' given the role of contextual-like sensitivity that emerges. One notable feature is that we do not get the right-left direction of standard T-biconditionals.)

11.7 Revision

Revision theorists share the basic revision-theoretic idea sketched above. Where revision theorists differ is in the details of how to extract *categorical* information from the various revision sequences. Such differences often show up as differences in how to define the revision operator at limit stages. For discussion of the main candidates and the issues that drive the differences, consult—in addition to sources cited in the main discussion (above)—the following works and their respective bibliographies.

- [Chapuis, 1996] offers so-called *fully varied* revision sequences that are contrasted with standard alternatives.
- [Chapuis and Gupta, 2000]. (See §11.1 for brief description.)
- [Shapiro, 2005], while not advancing a particular revision theory, contains a valuable examination of the philosophical import of revision rules.
- [Villanueva, 1997] contains essays directly on both philosophical and logical aspects of revision theory; it also contains essays on truth, in general.
- [Yaqūb, 1993] is a monograph advancing a distinctive set of revision rules.

11.8 Axiomatic Approaches

- [Cantini, 1996] is a monograph that provides a detailed discussion of axiomatic approaches to truth (and related notions).
- [Halbach and Horsten, 2002a] provides a user-friendly discussion of some of the main approaches (and surrounding issues) of axiomatic theories of truth.
- [Sheard, 1994] is a user-friendly discussion of axiomatic (and, to some extent, semantic) approaches to truth.

Acknowledgements

This essay is drawn from a forthcoming monograph. I am grateful to Hartry Field, Michael Glanzberg, Daniel Nolan, Graham Priest, and Lionel Shapiro for discussion and comments. I am also grateful to the following for discussion of various topics in the essay: Tim Bays, Marion David, Chris Gauker, Jay Garfield, Patrick Greenough, Anil Gupta, Carrie Jenkins, Michael Lynch, Doug Patterson, Stephen Read, Greg Restall, Kevin Sharp, Fritz Warfield, Crispin Wright, and also to audiences at Smith College, SUNY (Albany), University of Notre Dame, University of St Andrews, and members of the AHRC *Arché* Centre for Philosophy of Language, Logic, Mathematics, and Mind.

BIBLIOGRAPHY

[Aczel, 1988] Peter Aczel. *Non-Well-Founded Sets.* Number 14 in CSLI Lecture Notes. CSLI Publications, Stanford, 1988.

[Anderson and Belnap, 1975] Alan Ross Anderson and Nuel D. Belnap. *Entailment: The Logic of Relevance and Necessity*, volume 1. Princeton University Press, Princeton, 1975.

[Anderson *et al.*, 1992] Alan Ross Anderson, Nuel D. Belnap, and J. Michael Dunn. *Entailment: The Logic of Relevance and Necessity*, volume 2. Princeton University Press, Princeton, 1992.

[Armour-Garb and Beall, 2005] Brad Armour-Garb and JC Beall, editors. *Deflationary Truth.* Open Court Press, Chicago, 2005.

[Barwise and Etchemendy, 1987] Jon Barwise and John Etchemendy. *The Liar.* Oxford University Press, Oxford, 1987.

[Beall and Armour-Garb, 2005] JC Beall and Brad Armour-Garb, editors. *Deflationism and Paradox.* Oxford University Press, Oxford, 2005. In press.

[Beall and Restall, 2005] JC Beall and Greg Restall. *Logical Pluralism.* Oxford University Press, Oxford, 2005. In press.

[Beall *et al.*, 2005] JC Beall, Ross Brady, Alan Hazen, Graham Priest, and Greg Restall. Relevant restricted quantification. *Journal of Philosophical Logic*, 2005. Forthcoming.

[Beall, 2002] JC Beall. Deflationism and gaps: untying 'not's in the debate. *Analysis*, 62(4), 2002.

[Beall, 2003] JC Beall, editor. *Liars and Heaps: New Essays on Paradox.* Oxford University Press, Oxford, 2003.

[Beall, 2005a] JC Beall. Negation's holiday: double-aspect dialetheism. In David DeVidi and Tim Kenyon, editors, *A Logical Approach to Philosophy.* Kluwer Academic Publishers, Dordrecht, 2005. Forthcoming.

[Beall, 2005b] JC Beall. Transparent disquotationalism. In JC Beall and Brad Armour-Garb, editors, *Deflationism and Paradox*, pages 7–22. Oxford University Press, Oxford, 2005.

[Beall, 2006] JC Beall. True, false, and paranormal. *Analysis*, 2006. Forthcoming.

[Blackburn and Simmons, 1999] Simon Blackburn and Keith Simmons, editors. *Truth.* Oxford University Press, Oxford, 1999.

[Burge, 1979] Tyler Burge. Semantical paradox. *Journal of Philosophy*, 76:169–198, 1979. Reprinted in [Martin, 1984].

[Cantini, 1996] Andrea Cantini, editor. *Logical Frameworks for Truth and Abstraction: An Axiomatic Study.* Studies in Logic and the Foundations of Mathematics. Elsevier, Oxford, 1996.

[Chapuis and Gupta, 2000] André Chapuis and Anil Gupta, editors. *Circularity, Definition and Truth.* Indian Council of Philosophical Research, New Delhi, 2000.

[Chapuis, 1996] André Chapuis. Alternative revision theories of truth. *Journal of Philosophical Logic*, 25:399–423, 1996.

[Chellas, 1980] Brian F. Chellas. *Modal Logic: An Introduction.* Cambridge University Press, Cambridge, 1980.

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

[David, 1994] Marian David. *Correspondence and Disquotation.* Oxford University Press, Oxford, 1994.

[Dowden, 1984] Bradley Dowden. Accepting inconsistencies from the paradoxes. *Journal of Philosophical Logic*, 13:125–130, 1984.

[Dummett, 1978] Michael Dummett. Truth. In M. Dummett, editor, *Truth and Other Enigmas.* Duckworth, London, 1978.

[Feferman, 1984] Solomon Feferman. Toward useful type-free theories, I. *Journal of Symbolic Logic*, 49:75–111, 1984. Reprinted in [Martin, 1984].

[Field, 2001] Hartry Field. *Truth and the Absence of Fact*. Oxford University Press, Oxford, 2001.

[Field, 2002] Hartry Field. Saving the t-schema from paradox. *Journal of Philosophical Logic*, 31:1–27, 2002.

[Field, 2003a] Hartry Field. A revenge-immune solution to the semantic paradoxes. *Journal of Philosophical Logic*, 32:139–177, 2003.

[Field, 2003b] Hartry Field. The semantic paradoxes and the paradoxes of vagueness. In JC Beall, editor, *Liars and Heaps: New Essays on Paradox*, pages 262–311. Oxford University Press, Oxford, 2003.

[Field, 2004] Hartry Field. The consistency of the naïve (?) theory of properties. *The Philosophical Quarterly*, pages 78–104, 2004.

[Field, 2005a] Hartry Field. Is the liar both true and false? In JC Beall and Brad Armour-Garb, editors, *Deflationism and Paradox*, pages 23–40. Oxford University Press, Oxford, 2005. In press.

[Field, 2005b] Hartry Field. Solving the paradoxes, escaping revenge. 2005. To appear.

[Field, 2005c] Hartry Field. Variations on a theme by yablo. In JC Beall and Brad Armour-Garb, editors, *Deflationism and Paradox*, pages 53–74. Oxford University Press, Oxford, 2005. In press.

[Fine, 1975] Kit Fine. Vagueness, truth and logic. *Synthese*, 30:265–300, 1975.

[Fitting, 1986] Melvin Fitting. Notes on the mathematical aspects of Kripke's theory of truth. *Notre Dame Journal of Formal Logic*, 27:75–88, 1986.

[Gaifman, 1992] Haim Gaifman. Pointers to truth. *Journal of Philosophy*, 89:223–61, 1992.

[Glanzberg, 2001] Michael Glanzberg. The liar in context. *Philosophical Studies*, 103:217–251, 2001.

[Glanzberg, 2004a] Michael Glanzberg. A contextual–hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic*, 33:27–88, 2004.

[Glanzberg, 2004b] Michael Glanzberg. Quantification and realism. *Philosophy and Phenomenological Research*, 69:541–572, 2004.

[Glanzberg, 2004c] Michael Glanzberg. Truth, reflection, and hierarchies. *Synthese*, 142:289–315, 2004.

[Greenough and Lynch, 2005] Patrick Greenough and Michael P. Lynch. *Truth and Realism*. Oxford University Press, Oxford, 2005. In press.

[Grim, 1991] Patrick Grim. *The Incomplete Universe*. MIT Press, Cambridge, MA, 1991.

[Gupta and Belnap, 1993] Anil Gupta and Nuel Belnap. *The Revision Theory of Truth*. MIT Press, 1993.

[Gupta, 1982] Anil Gupta. Truth and paradox. *Journal of Philosophical Logic*, 11:1–60, 1982. Reprinted in [Martin, 1984].

[Hajek *et al.*, 2000] Petr Hajek, Jeff Paris, and John Shepherdson. The liar paradox and fuzzy logic. *Journal of Symbolic Logic*, 65:339–346, 2000.

[Halbach and Horsten, 2002a] Volker Halbach and Leon Horsten. Contemporary methods for investigating the concept of truth. In Volker Halbach and Leon Horsten, editors, *Principles of Truth*, Epistemische Studien. Ontos Verlag, 2002.

[Halbach and Horsten, 2002b] Volker Halbach and Leon Horsten, editors. *Principles of Truth*. Epistemische Studien. Ontos Verlag, 2002.

[Harman, 1986] Gilbert Harman. *Change In View: Principles of Reasoning*. Bradford Books. MIT Press, 1986.

[Herzberger, 1982] Hans Herzberger. Notes on naive semantics. *Journal of Philosophical Logic*, 11:61–102, 1982. Reprinted in [Martin, 1984].

[Hyde, 1997] Dominic Hyde. From heaps and gaps to heaps of gluts. *Mind*, 106:641–660, 1997.

[Jacquette, 2004] Dale Jacquette. Diagonalization in logic and mathematics. In Dov M. Gabbay and Franz Günthner, editors, *Handbook of Philosophical Logic*, pages 55–147. Kluwer Academic Publishers, Dordrecht, second edition, 2004.

[Kaplan, 1989] David Kaplan. Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes from Kaplan*, pages 481–614. Oxford University Press, New York, 1989.

[Kleene, 1952] S. C. Kleene. *Introduction to Metamathematics*. North-Holland, 1952.

[Koons, 1992] Robert C. Koons. *Paradoxes of Belief and Strategic Rationality.* Cambridge University Press, Cambridge, 1992.

[Kremer, 1988] Michael Kremer. Kripke and the logic of truth. *Journal of Philosophical Logic*, 17:225–278, 1988.

[Kremer, 2000] Philip Kremer. 'semantics' for languages with their own truth predicates. In André Chapuis and Anil Gupta, editors, *Circularity, Definition and Truth*, pages 217–46. Indian Council of Philosophical Research, New Delhi, 2000.

[Kripke, 1965] Saul A. Kripke. Semantical analysis of modal logic II: Non-normal modal propositional calculi. In Addison et al., editor, *The Theory of Models*, pages 206–220. North-Holland Publishing Co, 1965.

[Kripke, 1975] Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975. Reprinted in [Martin, 1984].

[Künne, 2003] Wolfgang Künne. *Conceptions of Truth.* Oxford University Press, New York, 2003.

[Lewis and Langford, 1932] C. I. Lewis and C. H. Langford. *Symbolic Logic.* The Century Co, New York and London, 1932.

[Lynch, 1998] Michael P. Lynch. *Truth in Context.* MIT Press, Cambridge, MA, 1998.

[Lynch, 2001] Michael P. Lynch, editor. *The Nature of Truth.* MIT Press, Cambridge, MA, 2001.

[Mares, 2004] Edwin D. Mares, editor. *Relevant Logic: A Philosophical Interpretation.* Cambridge University Press, Cambridge, 2004.

[Martin and Woodruff, 1975] Robert L. Martin and Peter W. Woodruff. On representing 'true-in-L' in L. *Philosophia*, 5:217–221, 1975. Reprinted in [Martin, 1984].

[Martin, 1970] Robert L. Martin, editor. *The Liar Paradox.* Yale University Press, New Haven, 1970.

[Martin, 1984] Robert L. Martin, editor. *Recent Essays on Truth and the Liar Paradox.* Oxford University Press, New York, 1984.

[Maudlin, 2004] Tim Maudlin. *Truth and Paradox: Solving the Riddles.* Oxford University Press, New York, 2004.

[McGee, 1991] Vann McGee. *Truth, Vagueness, and Paradox.* Hackett, Indianapolis, 1991.

[Meyer *et al.*, 1979] Robert K. Meyer, Richard Routley, and J. Michael Dunn. Curry's paradox. *Analysis*, 39:124–128, 1979.

[Norman *et al.*, 1989] Jean Norman, Graham Priest, and Richard Sylvan, editors. *Paraconsistent Logic: Essays on the Inconsistent.* Philosophia Verlag, 1989.

[Parsons, 1974a] Charles Parsons. The liar paradox. *Journal of Philosophical Logic*, 3:381–412, 1974. Reprinted in [Martin, 1984] and [Parsons, 1983].

[Parsons, 1974b] Charles Parsons. Sets and classes. *Noûs*, 8:1–12, 1974. Reprinted in [Parsons, 1983].

[Parsons, 1983] Charles Parsons. *Mathematics in Philosophy.* Cornell University Press, Ithica, 1983.

[Parsons, 1990] Terence Parsons. True contradictions. *Canadian Journal of Philosophy*, 20:335–353, 1990.

[Patterson, 2005] Douglas E. Patterson. Tarski, the liar, and inconsistent languages. *The Monist*, 89(1), 2005. In press.

[Priest *et al.*, 2004] Graham Priest, JC Beall, and Brad Armour-Garb, editors. *The Law of Non-Contradiction.* Oxford University Press, Oxford, 2004.

[Priest, 1979] Graham Priest. The logic of paradox. *Journal of Philosophical Logic*, 8:219–241, 1979.

[Priest, 1987] Graham Priest. *In Contradiction: A Study of the Transconsistent.* Martinus Nijhoff, The Hague, 1987.

[Priest, 1992] Graham Priest. What is a non-normal world? *Logique et Analyse*, 35:291–302, 1992.

[Priest, 1998] Graham Priest. What is so bad about contradictions? *Journal of Philosophy*, XCV(8):410–426, 1998. Reprinted in [Priest *et al.*, 2004].

[Priest, 2000] Graham Priest. Motivations for paraconsistency: The slippery slope from classical logic to dialetheism. In D. Batens, C. Mortensen, G. Priest, and J-P Van Bendegem, editors, *Frontiers of Paraconsistency.* Research Studies Press, KingÕs College Publications, Baldock, 2000.

[Priest, 2001] Graham Priest. *Beyond the Limits of Thought*. Oxford University Press, Oxford, second edition, 2001.

[Priest, 2002] Graham Priest. Paraconsistent logic. In Dov M. Gabbay and Franz Günthner, editors, *Handbook of Philosophical Logic (2nd Edition)*, volume 6, pages 287–393. Kluwer Academic Publishers, Dordrecht, 2002.

[Priest, 2005] Graham Priest. Spiking the Field artillery. In JC Beall and Brad Armour-Garb, editors, *Deflationism and Paradox*, pages 41–52. Oxford University Press, Oxford, 2005. In press.

[Priest, 2006a] Graham Priest. *Doubt Truth To Be A Liar*. Oxford University Press, Oxford, 2006. In press.

[Priest, 2006b] Graham Priest. *In Contradiction*. Oxford University Press, Oxford, second edition, 2006. Forthcoming.

[Quine, 1970] Willard van Orman Quine. *Philosophy of Logic*. Prentice-Hall, 1970.

[Ramsey, 1925] F. P. Ramsey. The foundations of mathematics. *Proceedings of the London Mathematical Society*, 25:338–384, 1925.

[Reinhardt, 1986] William N. Reinhardt. Some remarks on extending and interpreting theories with a partial predicate for truth. *Journal of Philosophical Logic*, 15:219–251, 1986.

[Rescher and Brandom, 1979] Nicholas Rescher and Robert Brandom. *The Logic of Inconsistency*. Rowman and Littlefield, Totowa, NJ, 1979.

[Restall, 1992] Greg Restall. Arithmetic and truth in Łukasiewicz's infinitely valued logic. *Logique et Analyse*, 139–140:303–312, 1992.

[Restall, 2000] Greg Restall. *An Introduction to Substructural Logics*. New York: Routledge, 2000.

[Routley and Loparic, 1978] Richard Routley and A. Loparic. Semantical analyses of Arruda-daCosta P-systems and adjacent non-replacement systems. *Studia Logica*, 37:301–320, 1978.

[Routley *et al.*, 1982] Richard Routley, Val Plumwood, Robert K. Meyer, and Ross T. Brady, editors. *Relevant Logics and their Rivals*. Ridgeview, 1982.

[Schantz, 2002] Richard Schantz, editor. *What is Truth?* Walter de Gruyter, New York, 2002.

[Shapiro, 2004] Stewart Shapiro. Simple truth, contradiction, and consistency. In Graham Priest, JC Beall, and Brad Armour-Garb, editors, *The Law of Non-Contradiction*. Oxford University Press, 2004.

[Shapiro, 2005] Lionel Shapiro. The rationale behind revision-rule semantics. *Philosophical Studies*, 2005. Forthcoming.

[Shaw-Kwei, 1954] Moh Shaw-Kwei. Logical paradoxes for many-valued systems. *Journal of Symbolic Logic*, 19:37–40, 1954.

[Sheard, 1994] Michael Sheard. A guide to truth predicates in the modern era. *Journal of Symbolic Logic*, 59(3):1032–54, 1994.

[Simmons, 1993] Keith Simmons. *Universality and The Liar*. Cambridge University Press, 1993.

[Skyrms, 1970] Brian Skyrms. Return of the liar: Three-valued logic and the concept of truth. *American Philosophical Quarterly*, 7:153–161, 1970.

[Skyrms, 1984] Brian Skyrms. Intensional aspects of semantical self-reference. In R. M. Martin, editor, *Recent Essays on Truth and the Liar*, pages 119–131. Oxford University Press, Oxford, 1984.

[Smullyan, 1992] Raymond M. Smullyan. *Gödel's Incompleteness Theorems*, volume 19 of *Oxford Logic Guides*. Oxford University Press, New York, 1992.

[Smullyan, 1993] Raymond M. Smullyan. *Recursion Theory for Metamathematics*, volume 20 of *Oxford Logic Guides*. Oxford University Press, New York, 1993.

[Soames, 1999] Scott Soames. *Understanding Truth*. Oxford University Press, New York, 1999.

[Tarski, 1936] Alfred Tarski. Der wahreitsbegriff in den formalisierten sprachen. *Studia Philosophica*, 1:261–405, 1936. Reprinted in English in [Tarski, 1956].

[Tarski, 1944] Alfred Tarski. The semantic concept of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4:341–376, 1944.

[Tarski, 1956] Alfred Tarski. *Logic, Semantics, Metamathematics: papers from 1923 to 1938*. Clarendon Press, Oxford, 1956. Translated by J. H. Woodger.

[van Fraassen, 1968] Bas C. van Fraassen. Presupposition, implication and self-reference. *Journal of Philosophy*, 65:136–52, 1968.

[van Fraassen, 1970] Bas C. van Fraassen. Truth and paradoxical consequences. In R. Martin, editor, *The Paradox of the Liar*, pages 13–23. Yale University Press, 1970.

[Varzi, 1999] Achille Varzi. *An Essay in Universal Semantics*, volume 1 of *Topoi Library*. Kluwer Academic Publishers, Boston, 1999.

[Villanueva, 1997] Enrique Villanueva, editor. *Truth*. Number 8 in Philosophical Issues. Ridgeview Publishing Company, Atascadero, California, 1997.

[Visser, 1984] Albert Visser. Four valued semantics and the liar. *Journal of Philosophical Logic*, 13:181–212, 1984.

[Visser, 2004] Albert Visser. Semantics and the liar paradox. In Dov M. Gabbay and Franz Günthner, editors, *Handbook of Philosophical Logic*, pages 149–240. Kluwer Academic Publishers, Dordrecht, second edition, 2004.

[White, 1979] R. B. White. The consistency of the axiom of comprehension in the infinite-valued predicate logic of łukasiewicz. *Journal of Philosophical Logic*, 8:509–534, 1979.

[Woodruff, 1984] P. Woodruff. Paradox, truth, and logic. *Journal of Philosophical Logic*, 13:213–252, 1984.

[Wright, 1987] Crispin Wright. *Realism, meaning and truth*. Oxford University Press, 1987.

[Wright, 1992] Crispin Wright. *Truth and Objectivity*. Harvard University Press, 1992.

[Yablo, 1993a] Stephen Yablo. Definitions, consistent and inconsistent. *Philosophical Studies*, 72:147–175, 1993. Special Issue on Definition.

[Yablo, 1993b] Stephen Yablo. Hop, skip, and jump. In James Tomberlin, editor, *Philosophical Perspectives: Language and Logic*, volume 7, pages 371–396. Ridgeview Publishing Co., Atascadero, CA, 1993.

[Yablo, 2003] Stephen Yablo. New grounds for naïve truth theory. In JC Beall, editor, *Liars and Heaps: New Essays on Paradox*, pages 313–330. Oxford University Press, Oxford, 2003.

[Yaqūb, 1993] Alladin Yaqūb. *The Liar Speaks the Truth*. Oxford University Press, 1993.