

Lying is a complicated matter, but what’s puzzling about sentences like the first one of this essay isn’t essentially tied to intentions, social norms, or anything like that. Rather, it seems to have something to do with truth, or at least, some semantic notion related to truth. The puzzle is usually named ‘the Liar paradox’, though this really names a family of paradoxes that are associated with our type of puzzling sentence. The family is aptly named one of paradoxes, as they seem to lead to incoherent conclusions, like that everything is true. Indeed, the Liar seems to allow us to reach such conclusions on the basis of logic, plus some very obvious principles that have themselves sometimes been counted as principles of logic. Thus, we have the rather surprising situation of something near or like logic leading us to incoherence. This is perhaps the most virulent strain of paradox, and suppressing it has been an important task in logic for about as long as there has been logic.

In this essay, we will review the important members of the family of Liar paradoxes, and some of the important ideas about how the paradoxes might be resolved. The past few thousand years have yielded a great number of proposals, and we will not be able to examine all of them. Instead, we will focus on a few that have have proved important in recent discussions.

1.1 Simple-falsity Liar

Consider a sentence named ‘FLiar’, which says of itself (i.e., says of FLiar) that it is false.

(FLiar) FLiar is false.

This seems to lead to contradiction as follows. If the sentence ‘FLiar is false’ is true, then FLiar is false. But if FLiar is false, then the sentence ‘FLiar is false’ is true. Since FLiar just is the sentence ‘FLiar is false’, we have it that FLiar is false if and only if FLiar is true. But, now, if every sentence is true or false, FLiar itself is either true or false, in which case—given our reasoning above—it is both true and false. This is a contradiction. Contradictions, according to many logical theories (e.g., classical logic, intuitionistic logic, and much more) imply absurdity—triviality, that is, that every sentence is true.

[RTG08], and, for instance, Read [Rea02, Rea06], Restall [Res08], and Simmons [Sim93].

An obvious response is to deny that every sentence is true or false, i.e. to deny the principle of bivalence. As we will discuss in §4, some descendants of this idea remain important in current work on the Liar. Even so, a simple variant Liar sentence shows that this immediate answer is not all there is to the story.

1.2 Simple-untruth Liar

Rather than work with falsehood, we can construct a Liar sentence with the complex predicate ‘not true’.² Consider a sentence named ‘ULiar’ (for ‘un-true’), which says of itself that it is not true.

(ULiar) ULiar is not true.

The argument towards contradiction is similar to the FLiar case. In short: if ULiar is true, then it is not true; and if not true, then true. But, now, if every sentence is true or not true, ULiar itself is true or not true, in which case it is both true and not true. This is a contradiction. According to many logical theories, a contradiction implies absurdity—triviality.

The two forms of the Liar paradox we have so far reviewed rely on some explicit self-reference—sentences talking directly about themselves. Such explicit self-reference can be avoided, as is shown by our next family of Liar paradoxes.

1.3 Liar cycles

Consider a very concise (viz., one-sentence-each) dialog between siblings Max and Agnes.

Max: Agnes’ claim is true.

Agnes: Max’s claim is not true.

What Max said is true if and only if what Agnes said is true. But what Agnes said (viz., ‘Max’s claim is not true’) is true if and only if what Max said is not true. Hence, what Max said is true if and only if what Max said is

²Terminology here is not uniform. Van Fraassen [vF68] introduced the term ‘strengthened Liar’ to name what we are calling the simple-untruth Liar. Van Fraassen’s term, however, has been more often used for a ‘revenge-like’ paradox in the manner of C. Parsons [Par74], as we discuss in §4.1.3.

not true. But, now, if what Max said is true or not true, then it is both true and not true. And this, as in the FLiar and ULiar cases, is a contradiction, implying, according to many logical theories, absurdity.

Liar paradoxes can also be formed using more complex sentence structure, rather than complex modes of reference. One that has been important involves Boolean compounds.

1.4 Boolean compounds

Boolean compounds can enter into Liar sentences in many ways. One relatively simple one is as follows. Consider the following sentence named ‘DLiar’ (for ‘Disjunctive’).

(DLiar) Either DLiar is not true or $1=0$.

Since $1 \neq 0$, we conclude (via Disjunctive Syllogism, which sanctions inferring sentence B from sentences $A \vee B$ and $\neg A$) that DLiar is not true. But a disjunction is true if and only if at least one disjunct is true. Hence, since ‘DLiar is not true’ is true, at least one of DLiar’s disjuncts is true (viz., the left one). Hence, DLiar is true. Hence, DLiar, like its cousins (FLiar, ULiar, etc), is true if and only if DLiar is not true. So, if DLiar is true or not true, it is both true and not true. This is a contradiction, which, according to many logical theories, implies absurdity.

We pause to mention DLiar as it is connected with another important paradox: Curry’s paradox. This involves conditionals that say of themselves only that if they (the conditional itself) are true, so too is some absurdity (e.g., ‘if this sentence is true, then $1=0$ ’ or ‘if this sentence is true, everything is true’ or so on). At least in languages where the conditional is the material conditional, and so $A \supset B$ is equivalent to $\neg A \vee B$, DLiar is equivalent to the Curry sentence ‘DLiar is true $\supset 1 = 0$ ’. Though this may set up some relations between the Liar and Curry’s paradox, we pause to note some important differences: the Curry paradox is most important where the conditional is more than the material conditional (or some modalized variant of it). In such settings, the Curry paradox does not wear negation on its sleeve, as DLiar does. For more on the Curry’s paradox, see [link: Curry’s Paradox].

1.5 Infinite sequences

The question of whether the Liar paradox really requires some sort of circularity has been the subject of extensive debate. Liar cycles (e.g., the Max–Agnes dialog) show that explicit self-reference is not necessary, but it is clear that such cycles involve circular reference. Yablo [Yab93b] has argued that a more complicated kind of multi-sentence paradox produces a Liar without circularity.

Yablo’s paradox relies on an infinite sequence of claims A_0, A_1, A_2, \dots , where each A_i says that all of the ‘greater’ A_k (i.e., the A_k such that $k > i$) are untrue. (In other words, each claim says of the rest that they’re all untrue.) Since we have an infinite sequence, this version of the Liar paradox appears to avoid the sort of circularity apparent in the previous examples; however, contradiction still seems to emerge. If A_0 is true, then all of the ‘greater’ A_k are untrue, and a fortiori A_1 is untrue. But, then, there is at least one true A_k greater than A_1 (i.e., some A_k such that $k > 1$), which contradicts A_0 . Conversely, if A_0 is untrue, then there’s at least one true A_k greater than A_0 . Letting A_m be such a one (i.e., a truth greater than A_0), we have it that A_{m+1} is untrue, in which case there’s some truth greater than A_{m+1} . But this contradicts A_m . What we have, then, is that if A_0 (the first claim in the infinite sequence) is true or untrue, then it is both. And this, as in the other cases, is a contradiction.

Whether Yablo’s paradox really avoids self-reference has been much-debated. See, for instance, Beall [Bea01], Priest [Pri97], Sorensen [Sor98], and also Cook [Coo06] and references therein.

2 Basic ingredients

We have already seen a kind of characteristic reasoning that goes with the Liar. We have also seen some common structure across all our example Liar paradoxes, such as the presence of truth predicates, and something like negation. We pause here to discuss these ingredients of the paradox, focusing on the *basic* Liars. Just what creates the Liar paradox, and just which of the puzzles we just surveyed is ‘basic’, is a contentious matter; different approaches to solving the Liar view these matters differently. Hence, our goal is merely to illuminate some common themes across different Liars, not to offer a full diagnosis of the source of the paradox.

We highlight three aspects of the Liar: the role of truth predicates, the kinds of principles for reasoning about truth that are needed, and the way that a paradox can be derived given these resources.

2.1 Truth predicate

The first ingredient in building a Liar is a truth predicate, which we write here as Tr . We follow the usual custom in logic of treating this as a predicate of sentences. However, especially as we come to consider some ways of resolving the Liar, it should be remembered that this treatment can be seen as more for convenience in exposition than a serious commitment to what truth bearers are.

We assume that we have, along with the truth predicate, appropriate names of sentences. For a given sentence A , suppose that $\ulcorner A \urcorner$ is a name for it. A predication of truth to A then looks like $Tr(\ulcorner A \urcorner)$.

We shall say that a predicate $Tr(x)$ is a *truth predicate for language \mathcal{L}* if $Tr(\ulcorner A \urcorner)$ is well-formed for every sentence A of \mathcal{L} . We typically expect Tr to obey some principles governing its behavior on sentences of a given language. It is to those we now turn.³

2.2 Principles of truth

The tradition, going back to Tarski [Tar35], is that the behavior of the truth predicate Tr is described by the following biconditional.

$$Tr(\ulcorner A \urcorner) \leftrightarrow A.$$

Indeed, Tarski took the biconditional here to be the material biconditional of classical logic. This is usually called the *T-schema*. For more on the T-schema, and Tarski's views of truth, see [link: Tarski, Alfred] and [link: Tarski, Alfred; Truth definitions].

³For the most part, when talking about a language, we mean an *interpreted language*, including a syntax, an interpretation or model, and valuation scheme which provides a consequence relation. Thus, we will often blur the distinction between a logic and a language. At some points, it will be a theory rather than a language that is at issue. We will note when these distinctions become important. But given the wide range of ideas in logic we will be surveying in rather brief form, we will often leave such subtleties to more leisurely presentations.

The Liar paradox has been a locus of thinking about non-classical logics (as we already saw a taste of, for instance, in the idea that bivalence might be rejected as part of a solution to the Liar). Thus, we should stop to consider what principles should govern the truth predicate Tr if classical logic is not to hold.

The leading idea for what might replace the T-schema points to two sorts of rules (e.g., two sorts of ‘inference rules’) that are characteristic of the truth predicate. If you have a sentence A , you can infer $Tr(\ulcorner A \urcorner)$, that is, you can ‘capture’ A with the truth predicate. Conversely, if you have $Tr(\ulcorner A \urcorner)$, you can infer A , that is, you can ‘release’ A from the truth predicate. The T-schema combines both the capture and release features, in classical logic. But it is often helpful to break these up, into:

- *Capture.* A implies $Tr(\ulcorner A \urcorner)$. (We also write this as $A \vdash Tr(\ulcorner A \urcorner)$.)
- *Release.* $Tr(\ulcorner A \urcorner)$ implies A . (We also write this as $Tr(\ulcorner A \urcorner) \vdash A$.)

‘Implies’ here is a logical notion. Just which one, and what the options are, depends on what background logic is assumed. For our discussion, we treat it as a ‘rule form’, so that if A implies B , the argument from A to B is valid. In some logical settings (e.g., classical logic, in which a certain so-called deduction theorem holds), this is equivalent to a conditional, in some, it is not. But even when they are not, capture and release jointly make A and $Tr(\ulcorner A \urcorner)$ logically equivalent in the sense of being inter-derivable. Thus, \vdash is being used here as a schematic placeholder for a range of different logical notions, each of which will provide some notion of valid inference in some logical theory.

.. *Parenthetical remark.* There are a number of logical subtleties here that we will not pursue, especially about how to formulate rules, and which rules are consistent. Different formulations of rules vary significantly in logical strength as well. See [link: truth, axiomatic theories of] for more on how consistent forms of capture and release can be formulated in classical logic. In the terminology of Friedman and Sheard [FS87], the rule forms of capture and release are called ‘T-Intro’ and ‘T-Elim’, and the conditional forms ‘T-In’ and ‘T-Out’. We prefer the broader terminology, since it highlights a general form of behavior common to a great variety of predicates and operators (e.g., *knowledge* releases but doesn’t capture; *possibility* captures but doesn’t release; and so on). *End parenthetical.* ..

2.3 The Liar in short

The Liar paradox begins with a language containing a truth predicate, which obeys some form of capture and release. We now explore more carefully how a paradox results from these assumptions.

2.3.1 Existence of Liar-like sentences

Putting aside Yablo-type paradoxes, the Liar relies on some form of self-reference, either direct, as in the simple Liars above, or indirect, as in Liar cycles. Most natural languages have little trouble generating self-reference. The first sentence of this essay is one example. Self-reference can be accidental, as in the case where someone writes ‘The only sentence on the blackboard in room 101 is not true’, by chance writing this in room 101 itself (as C. Parsons [Par74] noted).

In formal languages, self-reference is also very easy to come by. Any language capable of expressing some basic syntax can generate self-referential sentences via so-called diagonalization (or more properly, any language together with an appropriate theory of syntax or arithmetic). (See [link: logic-provability] and [link: Gödel-incompleteness theorems] for more details.) A language containing a truth predicate and this basic syntax will thus have a sentence L such that L implies $\neg Tr(\ulcorner L \urcorner)$ and vice versa:

$$L \dashv\vdash \neg Tr(\ulcorner L \urcorner).$$

This is a ‘fixed point’ of (the compound predicate) $\neg Tr$, and is, in effect, our simple-untruth Liar.

.. *Parenthetical remark.* It may help some readers (here and throughout) to think of such a Liar sentence L arising from a *name* c that denotes the sentence $\neg Tr(c)$. In this way, we can think of the existence of the Liar as being reflected in the identity $c = \ulcorner \neg Tr(c) \urcorner$. *End parenthetical.* ..

2.3.2 Other logical ‘laws’

Other conspicuous ingredients in common Liar paradoxes concern logical behavior of basic connectives or features of implication. A few of the relevant principles are:

- Excluded Middle (LEM): $\vdash A \vee \neg A$.

- Explosion (EFQ):⁴ $A, \neg A \vdash B$.
- Disjunction Principle (DP):⁵ If $A \vdash C$ and $B \vdash C$ then $A \vee B \vdash C$.

(This is not to suggest that these are the *only* logical features involved in common Liar paradoxes, but they're arguably the most important of the salient ones.)

2.3.3 The Liar in abstract

Given the foregoing ingredients, we can now give a slightly more abstract form of the paradox. (Our hope is to use this abstract form to highlight different responses to the paradox.) We suppose that we have a language \mathcal{L} with a truth predicate Tr , and that \mathcal{L} allows enough syntax to construct a sentence L such that $L \dashv\vdash \neg Tr(\ulcorner L \urcorner)$. We also suppose that the logic of \mathcal{L} enjoys LEM and EFQ and satisfies DP.

The argument that our Liar sentence L implies a contradiction runs as follows.

1. $Tr(\ulcorner L \urcorner) \vee \neg Tr(\ulcorner L \urcorner)$ [LEM]
2. Case One:
 - (a) $Tr(\ulcorner L \urcorner)$
 - (b) L [2a: release]
 - (c) $\neg Tr(\ulcorner L \urcorner)$ [2b: definition of L]
 - (d) $\neg Tr(\ulcorner L \urcorner) \wedge Tr(\ulcorner L \urcorner)$ [2a, 2c: adjunction]
3. Case Two
 - (a) $\neg Tr(\ulcorner L \urcorner)$
 - (b) L [3a: definition of L]
 - (c) $Tr(\ulcorner L \urcorner)$ [3b: capture]
 - (d) $\neg Tr(\ulcorner L \urcorner) \wedge Tr(\ulcorner L \urcorner)$ [3a, 3c: adjunction]
4. $\neg Tr(\ulcorner L \urcorner) \wedge Tr(\ulcorner L \urcorner)$ [1–3: DP]

⁴It is this principle/rule that we repeated above by saying that by the lights of many logical theories an arbitrary contradiction implies absurdity—triviality, all sentences.

⁵This is often called ‘ \vee -Out’ or ‘ \vee -Elim’ or, more suggestively, ‘Reasoning by Cases’.

This version of the Liar is one of many. With a little more complexity, for instance, either capture or release can be avoided in favor of some other background assumptions. Intuitionistic variants of the Liar are also available, though we shall not explore intuitionistic logic here.⁶

We have so far shown that with the given ingredients our Liar sentence L implies a contradiction (thus formalizing the reasoning in ULiar). From here, it is one short step to all-out absurdity—if the lone contradiction weren’t already absurd enough. We invoke EFQ to finish the proof. (Well, we also assume that $A \wedge B$ implies A and B , i.e., that Simplification is valid in \mathcal{L} .)

5. B [4: EFQ]

B , here, may be any—every—sentence that you like (or don’t like, as the case may be)! EFQ is the principle that every sentence follows from a contradiction; it sanctions the step from a single contradiction to outright triviality of logic.

In the face of such absurdity (triviality), we conclude that something is wrong in the foregoing Liar reasoning. The question is: what? This, in the end, is the question that the Liar paradox raises.

3 Significance

We have now seen that with some elementary assumptions about truth and logic, a logical disaster ensues. What is the wider significance of such a result?

From time to time, the Liar has been argued to show us something far-reaching about philosophy. For instance, Grim [Gri91] has argued that it shows the world to be essentially ‘incomplete’ in some sense, and that there can be no omniscient being. McGee [McG91] and others suggest that the Liar shows the notion of truth to be a vague notion. [link: vagueness] Glanzberg [Gla01] holds that the Liar shows us something important about the nature of context dependence in language, while Eklund [Ek102] holds that it shows us something important about the nature of semantic competence and the

⁶Again for the classical case, Friedman and Sheard [FS87] provide an exhaustive list of inconsistent theories relative to a fairly weak base theory. They show that either of the classical conditional forms of capture or release are inconsistent by themselves, over the base theory supplemented by the right choice of principles providing for completeness or consistency of truth.

languages we speak. Gupta and Belnap [GB93] claim that it reveals important properties of the general notion of definition. And there are other lessons, and variations on such lessons, that have been drawn.

Of more immediate concern, at least for our purposes in this essay, is what the Liar shows us about the basic principles governing truth, and about logic. In a skeptical vein, Tarski himself [Tar35, Tar44] seems to have thought the Liar shows the ordinary notion of truth to be incoherent, and in need of replacement with a more scientifically respectable one. (For more on Tarski, see [link: Tarski, Alfred] and [link: Tarski, Alfred-truth definitions]. For more on Tarski's aims and purposes, see Heck [Hec97].) More common, and perhaps the dominant thread in the solutions to the Liar, is the idea that the basic principles governing truth are more subtle than the T-schema reflects.

The Liar has also formed the core of arguments against classical logic, as it is some key features of classical logic that allow capture and release to result in absurdity. Most notable are the arguments for logics that are para-complete (e.g., Kripke [Kri75], Field [Fie08], and others) and paraconsistent (e.g., Asenjo [Ase66], Priest [Pri84, Pri06], and others).

In many cases inspired by wider views of the significance of the paradox, there have been a number of attempts to one way or another resolve the paradox. It is to these proposed solutions that we now turn.

4 Some families of solutions

In this section, we briefly survey some approaches to resolving the Liar paradox. We group proposed solutions into families, and try to explain the basic ideas behind them. In many cases, a full exposition would involve a great deal of technical material, that we will not go into here. Interested readers are encouraged to follow the references we provide for each basic idea.

4.1 Non-classical logics

One of the leading ideas for how to resolve the Liar paradox is that it shows us something about logic, in fact, something far-reaching about logic. The main idea is that the principles of capture and release are the fundamental conceptual principles governing truth, and cannot be modified. Instead, basic logic must be non-classical, to avoid a logical disaster of the kind we reviewed in §2.

One important way to motivate non-classical solutions is to appeal to a form of *deflationism* about truth. Such views take something akin to the T-schema to be the defining characteristic of truth, and as such, not open to modification. (See, e.g., Horwich [Hor90].) Most strictly, so-called transparency or ‘see-through’ or ‘pure disquotational’ conceptions of truth (e.g., Field [Fie94, Fie08] and Beall [Bea05, Bea09]) take the defining property of truth to be *intersubstitutability* of A and $Tr(\ulcorner A \urcorner)$ in all non-opaque contexts. This makes capture and release, in unrestricted form applying to all sentences of a language, a requirement for truth (at least where we have $A \vdash A$ or, more strongly, $\vdash A \rightarrow A$).⁷

Holding capture and release fixed, and applying it to all sentences without restriction, yields triviality unless logic is non-classical. There are two main sub-families of non-classical (transparency) truth theories: *paracomplete* and *paraconsistent*. We sketch the main ideas of each.

4.1.1 Paracomplete

According to paracomplete approaches to the Liar, the main lesson of the Liar is that LEM ‘fails’ in some sense. In other words: the Liar teaches us that some sentences (notably, Liars!) ‘neither hold nor do not hold’ (in some sense), and so are neither true nor false. As a result, the logic of truth is non-classical.

This idea is perhaps most natural in response to the simple-falsity Liar. There, it is tempting to say that there is some status other than truth and falsity, and the Liar sentence L has it. But this will not suffice, for instance, for the simple-untruth Liar. This says nothing about falsity. Rather, in some way the basic reasoning reviewed in §2.3 must fail, and the culprit, in the paracomplete view, is LEM. Liar-instances of LEM ‘fail’ (in some sense) according to the paracomplete approach; such sentences fall into the ‘gap’ between truth and falsity (to use a common metaphor).

There have been many proposals for using such non-classical logics to address the Liar. An early example is van Fraassen [vF68, vF70]. But Kripke’s work has been the most influential in recent times, not only to approaches to the Liar based on non-classical logic, but a range of other approaches we will survey in §4.2 as well. Thus, we pause to describe at least a little of Kripke’s framework.

⁷We have argued for a close connection between general views on the nature of truth and available avenues for resolving the paradox in our [BG08].

Kripke’s theory Logics where LEM fails are not themselves hard to come by. Among many such logics are a number of three-valued logics that allow sentences to take a third value over and above true and false. Sentences like Liar sentences take the third value. One of the most commonly applied logics is the Strong Kleene logic K_3 . We do not go into the details of K_3 here, but only note the properties of K_3 we need. (For more details, see [link: logic-many-valued], or Beall and van Fraassen [BvF03], or Priest [Pri08].) First and foremost, we have:

$$\not\vdash_{K_3} A \vee \neg A.$$

LEM fails. In fact, there are no logical truths (or valid sentences) according to K_3 . (We return to this on the topic of a ‘suitable conditional’ below.)

The challenge to using K_3 to flesh out a paracomplete theory is to explain how anything like (even rule-form) capture and release hold, and if you follow the deflationist line, how full unrestricted capture and release hold. One way of understanding the important work of Kripke [Kri75] (and related work of Martin and Woodruff [MW75]) is as a way of achieving just that.

Kripke begins with a fully classical language \mathcal{L}_0 containing no truth predicate (or more generally, no semantic terms). (Recall, we are assuming a language comes equipped with a consequence relation. For \mathcal{L}_0 it is classical.) He then considers extending it to a language \mathcal{L}_0^+ which contains a truth predicate Tr . The predicate Tr is taken to apply to every sentence of the expanded language \mathcal{L}_0^+ , including those of the original \mathcal{L}_0 . Thus, it is a self-applicative truth predicate (as the deflationist-inspired picture we mentioned must require), even though we begin with a language without a truth predicate.

We can think of \mathcal{L}_0 as interpreted by a classical model \mathcal{M}_0 . Kripke shows how to build an interpretation \mathcal{M}_0^+ for the expanded language. The main innovation is to see the truth predicate as *partial*. Rather than simply having an extension, it has an extension (set of things of which it is true), and an anti-extension (set of things of which it is false). The extension and anti-extension are mutually exclusive, but they need not jointly exhaust the domain of \mathcal{M}_0 . Pathological sentences like L fall in neither in the extension or the anti-extension of Tr .

Falling into neither the extension or the anti-extension of Tr acts like having a third value, and we can interpret \mathcal{L}_0^+ as acting like a language with a K_3 logic. Treating the language this way, Kripke shows up to build up a very plausible extension and anti-extension for Tr , typically written \mathcal{E} and

\mathcal{A} . The important property of the new extended model $\langle \mathcal{M}_0, \langle \mathcal{E}, \mathcal{A} \rangle \rangle$ is that the truth value of any sentence A and $Tr(\ulcorner A \urcorner)$ are exactly the same. A is true, false, or neither, just in case $Tr(\ulcorner A \urcorner)$ is. Furthermore, interpreting the expanded language \mathcal{L}_0^+ as a K_3 language, we have for K_3 consequence $A \dashv\vdash Tr(\ulcorner A \urcorner)$, just as we desired.

Kripke shows how to build up \mathcal{E} and \mathcal{A} by an inductive process. One starts with an ‘approximation’ of the extension and anti-extension of Tr , and successively improves it until the improvement process ceases to be productive (it reaches a ‘fixed point’). In fact, for the K_3 -based solution, the natural thing to do is start with an empty extension and anti-extension, and throw in sentences that are true at successive stages of the process.

Kripke’s construction can be applied to a number of different logics, including other many-valued logics such as the ‘Weak Kleene’ logic, and supervaluation logics. See, for instance, Burgess [Bur86] and McGee [McG91] for discussion. Kripke-style constructions engage a fair bit of mathematical subtlety. For an accessible overview of more of the details, see Soames [Soa99]. For a more mathematically rich exposition, see McGee [McG91].

Suitable conditionals Logics like K_3 suffer from the lack of a natural or ‘suitable’ conditional (in particular, one that satisfies $A, A \rightarrow B \vdash B$ and $\vdash A \rightarrow A$). This reveals a limitation of the Kripkean approach to the Liar. The language \mathcal{L}_0^+ cannot report the capture and release properties of truth itself in conditional form (i.e., T-biconditionals): Tr is transparent on this picture, and so $Tr(\ulcorner A \urcorner)$ and A are fully intersubstitutable. We don’t have $\neg A \vee A$ true for all sentences A in this theory, and hence don’t have $\neg Tr(\ulcorner A \urcorner) \vee A$ for all A . But $\neg Tr(\ulcorner A \urcorner) \vee A$ is equivalent to $Tr(\ulcorner A \urcorner) \rightarrow A$ in the theory, since (in the theory) \rightarrow is just the material conditional. The Kripke construction at hand, then, thus fails to enjoy all T-biconditionals—the natural candidates for expressing in the theory the basic capture and release features of truth.

A recent, major step towards supplementing Kripke’s framework with a suitable conditional is that of Field [Fie08]. Field’s theory is a major advance, but complex enough to be beyond the scope of this (very basic) introduction. Readers should consult Field’s own discussion for a taste of how such a modification might proceed. See Field [Fie08], and further discussion in Beall [Bea09].

4.1.2 Paraconsistent

Here, the basic idea is to allow the contradiction (e.g., up to and including step 4 of the derivation in §2.3.3), but alter the logic by rejecting EFQ—and, hence, avoid the absurdity involved in step 5.

Like the paracomplete approach we just surveyed, paraconsistent approaches to the Liar find easy, natural motivation in transparency or otherwise suitably ‘minimalist’ views of truth that require full intersubstitutability of A and $Tr(\ulcorner A \urcorner)$, and thus cannot restrict capture and release. But paraconsistent approaches have also found motivation in a Dummett-inspired anti-deflationist view, which takes the role of truth as the aim of assertion seriously (cf. Dummett [Dum59]). Indeed, Priest [Pri06] argues that this (non-transparency) view of truth motivates both the T-schema and LEM, and that this implies that the Liar sentence L is both true and not true. Hence, according to any such dialethic line (according to which at least one sentence is both true and not true), the only option is to reject EFQ.

Dialetheism Priest [Pri84, Pri06] has been one of the leading voices in advocating a paraconsistent approach to solving the Liar paradox. He has proposed a paraconsistent (and non-paracomplete) logic now known as LP (for *Logic of Paradox*), which retains LEM, but not EFQ.⁸ It has the distinctive feature of allowing true contradictions. This is what Priest calls the dialethic approach to truth. (See [SEP: Dialetheism] for more extensive discussion.)

Formally, LP can be seen as a three-valued logic; but where K_3 has truth-value gaps, LP has truth-value *gluts*. Thus, sentence in LP can be both true and false.

Likewise, Kripke-style techniques can be applied to produce an interpretation for a truth predicate, starting with a classical language \mathcal{L}_0 not containing a truth predicate. Again, an extension and anti-extension are assigned to Tr . Whereas Kripke’s original construction had the extension and anti-extension disjoint but not exhausting the domain, in this case we allow the extension and anti-extension to overlap, but suppose that the two together exhaust the domain of the model. Related techniques to Kripke’s can then be used to build an extension and anti-extension for Tr . The result

⁸The logic was first advanced by Asenjo [Ase66], though has come to be popularized by Priest’s work. The name ‘ LP ’ is due to Priest, while the term ‘dialetheism’ was coined by Priest and Routley.

is again an interpretation where A and $Tr(\ulcorner A \urcorner)$ get the same truth value in the model.

This construction was not given by Kripke himself, but variants have been pursued by a number of authors, including Dowden [Dow84], Priest [Pri84, Pri06], Visser [Vis84], and Woodruff [Woo84]. (A recent discussion of direct relevance is in Beall [Bea09].)

Combining paracompleteness and paraconsistency Though we have identified paracomplete and paraconsistent approaches to the Liar as two distinct options, they are not incompatible. Indeed, seen as theories of negation (if one wants), one might think that negation is neither exhaustive nor ‘explosive’ – i.e., satisfies neither LEM nor EFQ. An approach like this is the *FDE*-based (transparent) truth theory discussed in Dunn [Dun], Gupta and Belnap [GB93], Visser [Vis84], Woodruff [Woo84], Yablo [Yab93a], and—in effect—Brady [Bra89].

.. *Parenthetical remark.* The *LP*-based theories and K_3 -based theories are—at least on one (standard-first-order) level—simply strengthened logics of the broader *FDE* logic. For general discussion of such frameworks, see, e.g., Beall and van Fraassen [BvF03] and Priest [Pri08]. *End parenthetical.* ..

4.1.3 Expressive power and ‘revenge’

Working in classical logic, Tarski [Tar35] famously concluded from the Liar paradox that a language cannot define its own truth predicate. More generally, he took the lesson of the Liar to be that languages cannot express the full range of semantic concepts that describe their own workings. One of the main goals of the non-classical approaches to the Liar we have surveyed here is to avoid this conclusion, which many have seen as far too drastic. However, how successful these approaches have been in this regard remains a highly contentious issue.

In one sense, both the paracomplete and paraconsistent approaches achieve the desired result: they present languages which contain truth predicates which apply to sentences of that very language, and have the feature that A and $Tr(\ulcorner A \urcorner)$ have the same truth value. In this respect, they both present languages which contain their own truth predicate.

In the paracomplete case, the issue of whether this suffices has been much debated. The paracomplete view holds that the Liar sentence L is neither

true nor false, and this is key to retaining consistency. But note, the para-complete approach we discussed above cannot state this fact, as it cannot come out true that $\neg Tr(\ulcorner L \urcorner)$.

Just what to make of this has been debated. It is certainly the case that the set of true sentences in the kind of model Kripke constructs does not include $\neg Tr(\ulcorner L \urcorner)$. Because of this, some authors, such as McGee [McG91], T. Parsons [Par84], and Soames [Soa99] have in effect maintained that this is a further fact that goes beyond what the truth predicate needs to express, and so is immaterial to the success of the solution to the Liar. (Actually, McGee's view has another aspect, which we will discuss in §4.2.3.)

But nonetheless, it does appear that there is an important semantic fact about truth in the para-complete language, closely related to if not identical to a fact about truth per se, which the language cannot express. It thus has been argued to fail to achieve a fully adequate theory of truth. Kripke himself notes that there are some semantic concepts that cannot be expressed, and the argument has been pressed by Glanzberg [Gla04c] and C. Parsons [Par74].

One way of spelling out what is missing in the para-complete language is to introduce a new notion of *determinately*, so that the status of the Liar is that of not being determinately true. If so, then the Kripke para-complete language cannot express this concept of determinately. Some approaches taking para-complete ideas on board have sought to supplement the Kripke approach by adding notions of determinate truth. McGee [McG91] does so in a basically classical setting. In a non-classical, para-complete setting, Field [Fie08] supplements the basic para-complete approach with infinitely many different 'determinately' operators, each defined in terms of Field's 'suitable conditional', and each giving a different (stronger) notion of 'truth'. (See also some of the papers in Beall [Bea08].)

It is often argued in favor of paraconsistent approaches that they have no trouble 'characterizing' the status of Liars: they're true and false (i.e., true and have true negation). *LP* theories can state this. On the other hand, some such as Littmann and Simmons [LS04] and S. Shapiro [Sha04], have thought that there is a dual problem: namely, 'characterizing normal sentences'. (Some put this alleged problem as the problem of *just true*.) Whether this is a problem is something we leave open. (For some discussion, see Beall [Bea09], Field [Fie08], and Priest [Pri06].)

One other issue that arises here is that of so-called 'revenge paradoxes'. We can illustrate this with the simple-falsity Liar. Suppose one starts with this as the bench-mark Liar paradox, and proposes a simple solution that

rejects bivalence. In response one has shown the simple-untruth Liar, which undercuts the simple solution. This is the pattern of ‘revenge’, where a solution to the paradox is rejected on the basis of what might be taken to be a slightly modified form of the paradox. Revenge paradoxes for paracomplete solutions are often proposed: many points where the paracomplete language fails to express some semantic concept offer ways to construct a revenge problem. For instance, if the *determinately* avenue is taken, then one can construct a revenge problem via a sentence which says of itself that it is not determinately true.

We have seen at least some approaches (e.g. McGee [McG91] in some respects, T. Parsons [Par84], and Soames [Soa99]) reject the revenge problem, while some seek to solve it by additional apparatus (e.g. Field [Fie08]). As we discuss further in §4.3, contextualist views such as those of Burge [Bur79], Glanzberg [Gla01, Gla04a], and C. Parsons [Par74] tend to see revenge not as a separate problem, but as the core Liar phenomenon. For more discussion about revenge and its nature, see the papers in Beall [Bea08], and L. Shapiro [Shang].

4.2 Classical logic

We have now seen a range of options for responding to the Liar paradox by reconsidering basic logic. There are also a number of approaches that leave classical logic fixed, and try to find other ways of defusing the paradox.

One hallmark of most of these approaches is a willingness to somehow *restrict* the range of application of capture and release, to block the paradoxical reasoning. This is antithetical to the kind of deflationist view of truth we discussed in §4.1, but it is consistent with another view of truth. This other view takes the main feature of truth to be that it reports a non-trivial semantic property of sentences (e.g. corresponding with a fact in the world, or having a value in a model). Many approaches within classical logic embody the idea that a proper understanding of this feature allows for restricted forms of capture and release, and this in turn allows the paradox to be blocked, without any departure from classical logic.

We will consider a number of important approaches to the paradox within classical logic, most of which embody this idea in some form or another.

4.2.1 Tarski's hierarchy of languages

Traditionally, the main avenue for resolving the paradox within classical logic is Tarski's hierarchy of languages and metalanguages. Tarski concluded from the paradox that no language could contain its own truth predicate (in his terminology, no language can be 'semantically closed').

Instead, Tarski proposed that the truth predicate for a language is to be found only in an expanded metalanguage. For instance, one starts with an interpreted language \mathcal{L}_0 that contains no truth predicate. One then 'steps up' to an expanded language \mathcal{L}_1 , which contains a truth predicate, but one that only applies to sentences of \mathcal{L}_0 . With this restriction, it is easy enough to define a truth predicate which completely accurately states the truth values of every sentence in \mathcal{L}_0 , obeys capture and release, and yields no paradox. Of course, this process does not stop. If we want to describe truth in \mathcal{L}_1 , we need to step up to \mathcal{L}_2 to get a truth predicate for \mathcal{L}_1 . And so on. The process goes on indefinitely. At each stage, a new classical interpreted language is produced, which expresses truth for languages below it. (For more on the mathematics of this sort of hierarchy of languages, see Halbach [Hal97].)

Why is there no Liar paradox in this sort of hierarchy of languages? Because the restriction that no truth predicate can apply to sentences of its own language is enforced as a syntactic one. Any sentence L equivalent to $\neg Tr(\ulcorner L \urcorner)$ is not syntactically well-formed. There is no Liar paradox because there is no Liar sentence. See [link: Tarski, Alfred] and [link: Tarski, Alfred-truth definitions] for more on Tarski's views of truth.

Tarski's hierarchical approach has been subject to a number of criticisms. One is that in light of naturally occurring cases of self-reference, his ruling Liar sentences syntactically not well-formed seems overly drastic. Though Tarski himself was more concerned to resolve the Liar for formal languages, his solution seems implausible as applied to many naturally occurring uses of 'true'. Another important problem was highlighted by Kripke [Kri75]. As Kripke notes, any syntactically fixed set of levels will make it extremely hard, if not impossible, to place various non-paradoxical claims within the hierarchy. For instance, if Jc says *that everything Michael says is true*, the claim has to be made from a level of the hierarchy higher than everything Michael says. But, if among the things Micheal says is *that everything Jc says is true*, Michael's claims must be at a higher level than all of Jc 's claims. Thus, some of Michael's claims must be higher than some of Jc 's, and vice-versa. This is impossible.

In light of these sorts of problems, many have concluded that Tarski's hierarchy of languages and metalanguages buys a solution to the Liar paradox at the cost of implausible restrictiveness.

4.2.2 The closed-off Kripke construction

In light of these sorts of criticisms of Tarski's theory, a number of approaches to the Liar have sought to retain classical logic, but have some degree of self-applicability for the truth predicate. We know from the reasoning in §2.3 that some restrictions on capture and release will then be required. One goal has been to work out which ones are well-motivated, and how to implement them.

One way to do this was suggested by Kripke himself. Rather than see the Kripke apparatus we reviewed briefly in §4.1.1 as part of a non-classical logical approach, one can see it as an intermediate step towards building a classical interpretation of a self-applicative Tr .

Recall that the Kripke construction starts with a classical language \mathcal{L}_0 with no truth predicate. It passes to an expanded language \mathcal{L}_0^+ , but unlike a Tarskian metalanguage, this language contains a truth predicate Tr that applies to all of \mathcal{L}_0^+ . Kripke shows how to build a partial interpretation of Tr , providing an extension \mathcal{E} and an anti-extension \mathcal{A} . But one can then simply consider the *classical* model $\langle \mathcal{M}_0, \mathcal{E} \rangle$, using only the extension. This is the 'closed-off' construction, as the gap between extension and anti-extension is closed off by throwing everything in the gap into the false category of a classical model.

We know this interpretation cannot make true all of capture and release. But it does make a restricted form true. The following holds in the closed-off model:

$$[Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner)] \rightarrow [Tr(\ulcorner A \urcorner) \leftrightarrow A].$$

This tells us that capture and release (in the form of the T-schema) holds for sentences that are well-behaved, in the sense of satisfying $Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner)$.

What happens to the Liar sentence on this approach? As in the three-valued case, the Liar is interpreted as falling within the gap. L is neither in \mathcal{E} nor \mathcal{A} . L thus falls outside of the domain where Tr is interpreted as well-behaved. Because the situation is classical, and $\ulcorner L \urcorner \notin \mathcal{E}$, we know that $\neg Tr(\ulcorner L \urcorner)$ is true in the closed-off model; likewise, so is $\neg Tr(\ulcorner \neg L \urcorner)$.

On well-behaved sentences, we have the fixed point property that A and $Tr(\ulcorner A \urcorner)$ have the same truth value, and so the semantics of \mathcal{L}_0^+ and the semantics it assigns to Tr correspond exactly. On pathological sentences like L , they do not, and indeed, cannot, on pain of triviality.

In a point related to the closed-off construction, it was observed by Feferman [Fef84] that if we are careful about negation, we can dispense with \mathcal{A} altogether in the Kripke construction. Thus, the construction can be done without any implicit appeal to many-valued logic. Related ways of thinking about Kripke’s construction are discussed McGee [McG91], and are applied in Glanzberg [Gla04a].

4.2.3 Determinateness revisited

In §4.1.3 we noted that paracomplete approaches to the paradox can be vulnerable to ‘revenge paradoxes’ based on some idea of indeterminate truth or lacking a truth value. Related issues bear in the classical case. We will discuss a few in turn.

Grounding The closed-off Kripke construction can help fill in the idea of a *determinately* operator discussed in §4.1.3. Instead of an operator, it allows us to define a predicate $D(\ulcorner A \urcorner)$ by $Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner)$. D represents ‘determinately’ in the sense of applying to sentences that have a truth value according to Tr , as it were, ‘determined’ by the model produced by the Kripke construction. It also, as we observed, applies to all the sentences which are well-behaved in the sense of obeying the T-schema (or capture and release).

Formally, the sentences to which D applies in the model generated by the Kripke construction are those which fall in \mathcal{E} or have their negations fall in \mathcal{E} (equivalently fall in \mathcal{A}). Kripke labeled this being *grounded*.⁹

It has often been noted that there is also a more informal notion of determinateness or grounding, to which the formal notion expressed by D at least roughly corresponds (cf. Herzberger [Her70]). The idea is that the determinate sentences are the ones with well-defined semantic properties. Where we have no such well-defined semantic properties, we should not expect the

⁹A technical note: Kripke defined being grounded as appearing in the *minimal* fixed point. We have not discussed other fixed points here, but there are many, and a sentence can appear in them without appearing in the minimal one.

truth predicate to report anything well-behaved, nor should we expect properties like Capture and Release to hold. Kripke’s construction builds up \mathcal{E} in stages, starting with sentence with no semantic terms, and adding semantic complexity at each stage. One reaches \mathcal{E} at the limit of this process, which allows us to think of \mathcal{E} as indicating the limit of where semantic values are assigned by a well-defined process. Thus, the formal notion of grounding provided by D is sometimes suggested to reflect the extent to which sentences have well-defined semantic properties.

McGee on truth and definite truth Another view which makes use of a form of determinateness is advocated by McGee [McG91]. McGee’s theory, like many we have surveyed, here, is rich in complexity to which we cannot do justice. The theory has many components, including a mathematically sophisticated approaches to truth related to the Kripkean ideas we have been discussing, in a setting which holds to classical logic.

McGee relies on two notions: truth and definite truth. Definite truth is a form of the idea we glossed as determinateness. Formally, for McGee, definite truth is identified with provability in a partially interpreted language, using an extension of classical logic which takes in facts about the partial interpretation known as \mathcal{A} -logic. It is thus different from the grounding notion we just discussed. McGee treats definitely as a *predicate*, on par with the truth predicate, and not as an operator on sentences as some developments do. With the right notion of definite truth, McGee shows that a partially interpreted language containing its own truth predicate can meet restricted forms of capture and release put in terms of definite truth. Where Def is the definiteness predicate, McGee show how to link truth and definite truth, by showing how to validate:

$$\begin{aligned} Def(\ulcorner A \urcorner) &\text{ iff } Def(\ulcorner Tr(\ulcorner A \urcorner) \urcorner). \\ Def(\ulcorner \neg A \urcorner) &\text{ iff } Def(\ulcorner \neg Tr(\ulcorner A \urcorner) \urcorner). \end{aligned}$$

Indeed, McGee shows that these conditions can be met within a theory of both truth and definite truth, where truth meets appropriate forms of capture and release, and also where a formal statement of bivalence for truth comes out definitely true. McGee thus provides a theory which has strongly self-applicative truth and definite truth, within a classical setting.

Though truth may satisfy the formal property of bivalence, it is crucial to McGee’s approach that definite truth is an open-ended notion, which may be

strengthened (formally, by strengthening a partially interpreted language). Thus, definite truth meets weaker forms of capture and release than truth itself. (Some instances of $Def(\ulcorner A \urcorner) \rightarrow A$ fail to be definitely true, according to McGee.) Furthermore, McGee suggests that this behavior of truth and definite truth makes truth a *vague* predicate. It remains disputed whether McGee’s theory avoids the kind of revenge problems that plague other Kripkean approaches. An argument that it is vulnerable to forms of revenge is given in Glanzberg [Gla04c].

4.2.4 Other classical approaches

We have now surveyed some important representatives of approaches to resolving the Liar within classical logic. There are a number of others, many of them involving some complex mathematics. We will pause to mention a few of the more important of these, though given the mathematical complexity, we will only gesture towards them.

Axiomatic theories of truth There is an important strand of work in proof theory, which has sought to develop axiomatic theories of self-applicative truth in classical logic, including work of Feferman [Fef84, Fef91], Cantini [Can96], and Friedman and Sheard [FS87]. The idea is to find ways of expressing rules like capture and release that retain consistency. Options include more care about how proof-theoretic rules of inference are formulated, and more care about formulating restricted rules. The main ideas are discussed in [link: truth-axiomatic theories of], to which we will leave the details.

Truth and inductive definitions Kripke’s work on truth was developed in conjunction with some important ideas about inductive definitions (as we see, for instance, in the later parts of Kripke [Kri75]). These connections are explored further in work of Burgess [Bur86] and McGee [McG91]. We also pause to mention work of Aczel [Acz80] combining ideas about inductive definitions and the lambda calculus.

4.3 Contextualist approaches

Another family of proposed solutions to the Liar are *contextualist solutions*. These also rely make use of classical logic, but base their solutions primarily

on some ideas from the philosophy of language. They take the basic lesson of the Liar to be that truth predicates show some form of *context dependence*, even in otherwise non-context-dependent fragments of a language. They seek to explain how this can be so, and rely on it to resolve the problems faced by the Liar.

Contextualist theories share with a number of approaches we have already seen the idea that there is something indeterminate or semantically not well-formed about our Liar sentence L . But, contextualist views give a special role to issues of ‘revenge’ and lack of expressive power.

4.3.1 Instability and revenge

One way of thinking about why the truth predicate is not well-behaved on the Liar sentence is that there is not really a well-defined truth bearer provided by the Liar sentence. To make this vivid (as discussed by C. Parsons [Par74] and subsequently Glanzberg [Gla01]), suppose that truth bearers are propositions expressed by sentences in contexts, and that the Liar sentence fails to express a proposition. This is the beginnings of an account of how the Liar winds up ungrounded or in some sense indeterminate. At least, we should not expect Tr to be well-behaved where sentences fail to express propositions.

But, it is an unstable proposal. We can reason that if the Liar sentence fails to express a proposition, it fails to express a true proposition. In the manner of a revenge paradox, if our Liar sentence had originally said ‘this sentence does not express a true proposition’, then we would have our Liar sentence back. And, we have shown that this sentence says something true, and so expresses a true proposition. Thus, from the assumption that the Liar sentence is indeterminate or lacks semantic status, we reason that it must have proper semantic status, and indeed say something true. We are hence back in paradox.

Contextualists do not see this as a new ‘revenge’ paradox, but the basic problem posed by the Liar. First of all, in a setting where sentences are context dependent, the natural formulation of a truth claim is always in terms of expressing a true proposition, or some related semantically careful application of the truth predicate. But more importantly, to the contextualist, the main issue behind the Liar is embodied in the reasoning on display here. It involves two key steps. First, assigning the Liar semantically defective status—failing to express a proposition or being somehow indeterminate. Second, concluding from the first step that the Liar must true—and so not

indeterminate or failing to express a proposition—after all. Both steps appear to be the result of sound reasoning, and so both must be true. The main problem of the Liar, according to a contextualist, is to explain how this can be, and how the second step can be non-paradoxical. (Such reasoning is explored by Glanzberg [Gla01, Gla04c] and C. Parsons [Par74]. For a critical discussion, see Gauker [Gau06].)

Thus, contextualists seek to explain how the Liar sentence can have unstable semantic status, switching from defective to non-defective in the course of this sort of inference. They do so by appealing to the role of *context* in fixing the semantic status of sentences. Sentences can have different semantic status in different contexts. Thus, to contextualists, there must be some non-trivial effect of context involved in the Liar sentence, and more generally, in predication of truth.

4.3.2 Contextual parameters on truth predicates

One prominent contextualist approach, advocated by Burge [Bur79] and developed by Koons [Koo92] and Simmons [Sim93], starts with the idea that the Tarskian hierarchy itself offers a way to see the truth predicate as context dependent. Tarski’s hierarchy postulates a hierarchy of truth predicates Tr_i . What if i is not merely a marker of level in a hierarchy, but a genuine contextual parameter? If so, then the Liar sentence is in fact context-dependent: it has the form $\neg Tr_i(\ulcorner L \urcorner)$, where i is set by context. Context then sets the level of the truth predicate.

This idea can be seen as an improvement on the original Tarskian approach in several respects. First, once we have a contextual parameter, the need to insist that Liar sentences are never well-formed disappears. Hence, we can think of each Tr_i as including some limited range of applicability to sentences its own language. Using the Kripkean techniques like the closed-off construction we reviewed above, predicates like Tr_i can be constructed which have as much self-applicability as Kripke’s own. (Burge [Bur79] and the postscript to C. Parsons [Par74] consider briefly how Kripke techniques could be applied in this setting. Though he works in a very different setting, ideas of Gaifman [Gai88, Gai92] can be construed as showing how even more subtle ways of interpreting a context-dependent truth predicate can be developed.)

With suitable care, other problems for the Tarskian hierarchy can be avoided as well. Burge proposes that the parameter i in Tr_i is set by a

Gricean pragmatic process. In effect, speakers implicate that i is to be set to a level for which the discourse they are in can be coherently interpreted (with a maximal coherent extension for Tr_i). Thus, truth does indeed find its own level, and so Kripke’s objection about how to fix levels for non-paradoxical sentences may be countered.

This approach gives substance to the idea that the Liar sentence is context dependent. Any sentence containing Tr_i will be context dependent, inheriting a contextual parameter along the way. This offers a way to make sense of the arguments for the instability of the semantic status of L that motivated contextualism. In an initial context, we fix some level i . This is the level at which L is interpreted. Call this interpretation L_i . L_i says $\neg Tr_i(\ulcorner L_i \urcorner)$. By the usual Liar reasoning, we show that L_i must lack determinate semantic status—or fail to express a proposition. As we discussed, we then reason that L must come out true. According to the contextualist view at hand, this is the claim that L_i is true according to some other context, where a wider truth predicate is in play. This amounts to being true at some higher level of the hierarchy. We can conclude, for instance, that the Liar sentence as it was used at level i is true according to a wider level $k > i$. Hence, $Tr_k(\ulcorner L_i \urcorner)$, where $k > i$.

This form of contextualism thus maintains that once we see the context-dependent behavior of Tr_i , we can make good sense of the instability of L . This can be seen as an improvement on both the Tarskian view, and embodying some of the techniques of classical logic we reviewed in §4.2. Depending how the Burge view is spelled out technically, it will either have full capture and release at each level, or capture and release with the same restrictions as the closed-off Kripke construction.

The view that posits contextual parameters on the truth predicate does face a number of questions. For instance, it is fair to ask why we think the truth predicate really has a contextual parameter, especially if we mean a truth predicate like the one we use in natural language. Merely noting that such a parameter would avoid paradox does not show that it is present in natural language. Furthermore, whether it is acceptable to see truth as coming in levels at all, context-based or not, remains disputed. (Not all those who advocate contextual parameters on the truth predicate agree about the role of hierarchy. In particular, Simmons [Sim93] advocates a view he labels the ‘singularity theory’ which he proposes avoids outright hierarchical structures.) Finally, the Burgean appeal to Gricean mechanisms to set levels of truth has been challenged. (For instance, Gaifman [Gai92] asks if the

Gricean process does any substantial work in Burge’s account.)

Contextualist approaches come in many varieties, each of which makes use of slightly different apparatus. With contextualist theories the choice often turns on issues in philosophy of language as well as logic. We already noted a different way of developing contextualist ideas from Gaifman [Gai88, Gai92]. We will now briefly review a few more alternatives.

4.3.3 Contextual effects on quantifier domains

Another contextualist approach, stemming from work of C. Parsons [Par74] and developed by Glanzberg [Gla01, Gla04a], seeks to build up the context dependence of the Liar sentence, and ultimately the context dependence of the truth predicate, from more basic components. The key is to see the context dependence of the Liar sentence as derived from the context dependence of quantifier domains.

Quantification enters the picture when we think about how to account for predication of truth when sentences display context dependence. In such an environment, it does not make good sense to predicate truth of sentences directly. Not all sentences will have the right kind of determinate semantic properties to be truth bearers; or, as we have been putting it, not all sentences will express propositions. But then, to say that a sentence S is true in context c is to say that *there is* a proposition p expressed by S in c , and that proposition p is true.

The current contextualist proposal starts with the observation that quantifiers in natural language typically have context-dependent domains of quantification. When we say ‘Everyone is here’, we do not mean everyone in the world, but everyone in some contextually provided subdomain. Context dependence enters the Liar, according to this contextualist view, in the contextual effects on the domain of the propositional quantifier $\exists p$.

In particular, this domain must *expand* in the course of the reasoning about the semantic status of the Liar. In the initial context, $\exists p$ must range over a small enough domain that there is no proposition for L to express. In the subsequent context, the domain expands to allow L to express some true proposition. Proposals for how this expansion happens, and how to model the truth predicate and the relation of expressing a proposition in the presence of the Liar, have been explored by Glanzberg [Gla01, Gla04a]), building on work of C. Parsons [Par74]. Defenders of this approach argue that it does better in locating the locus of context dependence than the parameters on

truth predicates view.

4.3.4 Situation theory

Another variant on the contextualist strategy for resolving the Liar, developed by Barwise and Etchemendy [BE87] and Groeneveld [Gro94], relies on *situation theory* rather than quantifier domains to provide the locus of context dependence. Situation theory is a highly developed part of philosophy of language, so we shall again give only the roughest sketch of how their view works.

A situation is a partial state the world might be in: something like a being F . Situations are classified by what are called situation types. A proposition involves classifying a situation as being of a situation type. Thus, a proposition $\{s; [\sigma]\}$ tell us that situation s is of type σ . The situation s here plays a number of roles, including that of providing a context.

When it comes to the Liar, Barwise and Etchemendy construe Liar propositions as having the form $f_s = \{s; [Tr, f_s; 0]\}$, relative to an initial situation s . This is a proposition f_s which says of itself that its falsity is a fact that holds in s . There is a sense in which this proposition cannot be expressed. In particular, the state of affairs $\langle Tr, f_s; 0 \rangle$ cannot be in s . (Actually, Barwise and Etchemendy say that the proposition is expressible, but give up on what they call the F -closure of s . But there is a core observation in common between these two points, and the details do not matter for our purposes here.) There is then a distinct situation $s' = s \cup \{\langle Tr, f_s; 0 \rangle\}$, and the proposition $\{s'; [Tr, f_s; 0]\}$ relative to this new situation—this new ‘context’—is true.

This idea clearly has a lot in common with the restriction on quantifier domains view. In particular, both approaches seek to show how the domain of contents expressible in contexts can expand, to account for the instability of the Liar sentence. For discussion of relations between the situation-theoretic and quantifier domain approaches, see Glanzberg [Gla04a]. Barwise and Etchemendy discuss relations between their situation-based and a more traditional approach in [BE87, Ch. 11]. For a detailed match-up between the Barwise and Etchemendy framework and a Burgean framework of indexed truth predicates, see Koons [Koo92].

4.3.5 Issues for contextualism

It is a key challenge to contextualists to provide a full and well-motivated account of the source and nature of the shift in context involved in the Liar, though of course, many contextualists believe they have met this challenge. In favor of the contextualist approach is that it takes the revenge phenomenon to be the basic problem, and so is largely immune to the kinds of revenge issues that affect other approaches we have considered. But, it may be that there is another form of revenge which might be applied. To retain consistency, contextualists must apply restrictions on quantifiers to such quantifiers as ‘all contexts’. To achieve this, it must presumably be denied that there are any absolutely unrestricted quantifiers. Glanzberg [Gla04b, Gla06] argues this is the correct conclusion, but it is highly controversial. For a survey of thinking about this, see the papers in Rayo and Uzquiano [RU06].

4.4 The revision theory

Another approach to the Liar, advocated by Gupta [Gup82], Herzberger [Her82], Gupta and Belnap [GB93], and a number of others, is the *revision theory of truth*. This approach shares some features with the views we surveyed in §4.2, in that it takes classical logic for granted. We also believe it has an affinity with the views discussed in §4.3, as it rethinks some basic aspects of semantics. But it is a distinctive approach. We will sketch some of the fundamentals of this view. For a discussion of the foundations of the revision theory, and its relations to contextualism, see L. Shapiro [Sha06]. More details, and more references, may be found in [link: Truth-revision theory of].

The revision theory of truth starts with the idea that we may take the T-schema at face value. Indeed, Gupta and Belnap [GB93] take up a suggestion from Tarski [Tar44], that the instances of the T-schema can be seen as partial definitions of truth; presumably all the instances together, for the right language or family of languages constituting a complete definition. At the same time, the revision theory holds fast to classical logic. Thus, we already know, we have the Liar paradox for any language with enough expressive resources to produce Liar sentences.

In response, the revision theory proposes a different way of approaching the semantic properties of the truth predicate. In keeping with our practices here, we may begin with a classical model \mathcal{M}_0 for a language \mathcal{L}_0 without a

truth predicate, and consider what happens when we add a truth predicate Tr to form the extended language \mathcal{L}_0^+ . This language has a full self-applicative truth predicate, and so can generate the Liar sentence L .

To build a classical model for \mathcal{L}_0^+ , we need an extension for Tr . Let us pick a set: call it H for a *hypothesis* about what the extension of Tr might be. H may be \emptyset , it may be the entire domain of \mathcal{M}_0 , or it may be anything else. It need not be a particularly good approximation of the semantic properties of Tr .

Even if it is not, $\langle \mathcal{M}_0, H \rangle$ still provides a classical model, in which we can interpret \mathcal{L}_0^+ . With that, we can in effect apply the T-schema, relative to our hypothesis H , and see what we get. More precisely, we can let $\tau(H) = \{ \ulcorner A \urcorner : \ulcorner A \urcorner \text{ true in } \langle \mathcal{M}_0, H \rangle \}$. $\tau(H)$ is generally a better hypothesis about what is true in our language than H might have been. At least, clearly, if H made foolish guesses about the truth of sentence of the truth-free fragment \mathcal{L}_0 , they are corrected in $\tau(H)$, which contains everything from \mathcal{L}_0 true in \mathcal{M}_0 . Thus, $\langle \mathcal{M}_0, \tau(H) \rangle$ is generally a better model of \mathcal{L}_0^+ than $\langle \mathcal{M}_0, H \rangle$.

Better in many respects. But when it comes to paradoxical sentences like L , we see something different. As a starting hypothesis, let us consider $H = \emptyset$. Consider what happens to the truth of L as we apply τ :

n	truth value of L in $\langle \mathcal{M}_0, \tau^n(\emptyset) \rangle$
0	true
1	false
2	true
3	false
4	true
\vdots	\vdots

The Liar sentence never stabilizes under this process. We reach an alternation of truth values which will go on for ever.

In the terminology of the revision theory, τ is a *revision rule*. It takes us from one hypothesis about the interpretation of Tr to another. Sequences of values we generate by such revision rules, starting with a given initial hypothesis, are *revision sequences*. We leave to a more full presentation the important issue of the right way to define *transfinite* revision sequences. (See [link: Truth-revision theory of].)

The characteristic property of paradoxical sentences like the Liar sentence is that they are unstable in revision sequences: there is no point in the

sequence at which they reach a stable truth value. This classifies sentences as stably true, stably false, and unstable. The revision theory develops notions of consequence based on these, and related notions. See [link: Truth-revision theory of] for further exposition of this rich theory.

4.5 Inconsistency views

In §2.3.3 we saw that the Liar paradox, in the presence of unrestricted capture and release and classical logic, leads to contradiction. So long as we have EFQ (as classical logic does), this results in triviality. Most of the proposed solutions we have considered (with the exception of the revision theory) try to avoid this result somehow, either by restricting capture and release or departing from classical logic. But there is another idea that has occasionally been argued, that the Liar paradox simply shows that the kinds of languages we speak, which contain their own truth predicates, are *inconsistent*.

This is not an easy view to formulate. Though Tarski himself seemed to suggest something along these lines (for natural languages, specifically), it was argued by Herzberger [Her67] that it is impossible to have an inconsistent language.

In contrast, Eklund [Ekl02] takes seriously the idea that our semantic intuitions, expressed, for instance, by unrestricted capture and release, really are inconsistent. Eklund grants that this does not make sense if these intuitions have their source simply in our grasp of the truth conditions of sentences. But he suggests an alternative picture of semantic competence which does make sense of it (closely related to conceptual role views of meaning). He suggests that we think of semantic competence in terms of a range of principles speakers are disposed to accept in virtue of knowing a language. Those principles may be inconsistent. But even so, they determine semantic values. Semantic values will be whatever comes closest to satisfying the principles—whatever makes them maximally correct—even if nothing can satisfy all of them due to an underlying inconsistency.

Eklund thus supports an idea suggested by Chihara [Chi79]. Chihara's main aim is to provide what he calls a *diagnosis* of the paradox, which should explain why the paradox arises and why it appears compelling. But along the way, he suggests that the source of the paradox is our acceptance of the T-schema (by convention, he suggests), in spite of its inconsistency.

A related, though distinct, view is defended by Patterson [Pat07, Patng]. Patterson argues that competence with a language puts one in a cognitive

state to an inconsistent theory—one including the unrestricted T-schema and governed by classical logic. He goes on to explore how such a cognitive state could allow us to successfully communicate, in spite of relating us to a false theory.

5 A few concluding remarks

There is much more to say about the Liar paradox than we have covered here. There are more approaches to the Liar variants we have mentioned, and more related paradoxes like those of donation, properties, etc. There are also more important technical results, and more important philosophical implications and applications. Our goal here has been to be more suggestive than exhaustive, and we hope to have given the reader an indication of what the Liar paradox is, and what its consequences might be.

References

- [Acz80] Peter Aczel. Frege structures and the notions of proposition, truth and set. In J. Barwise, H. J. Keisler, and K. Kunen, editors, *The Kleene Symposium*, pages 31–59. North-Holland, Amsterdam, 1980.
- [And70] Alan Ross Anderson. St. Paul’s epistle to Titus. In Robert L. Martin, editor, *The Paradox of the Liar*, pages 1–11. Ridgeview, Atascadero, 1970.
- [Ase66] F. G. Asenjo. A calculus of antinomies. *Notre Dame Journal of Formal Logic*, 16:103–105, 1966.
- [BE87] Jon Barwise and John Etchemendy. *The Liar*. Oxford University Press, Oxford, 1987.
- [Bea01] Jc Beall. Is Yablo’s paradox non-circular? *Analysis*, 61:176–187, 2001.
- [Bea05] Jc Beall. Transparent disquotationalism. In Jc Beall and Bradley Armour-Garb, editors, *Deflationism and Paradox*, pages 7–22. Oxford University Press, Oxford, 2005.

- [Bea08] Jc Beall, editor. *Revenge of the Liar*. Oxford University Press, Oxford, 2008.
- [Bea09] Jc Beall. *Spandrels of Truth*. Oxford University Press, Oxford, 2009.
- [BG08] Jc Beall and Michael Glanzberg. Where the paths meet: Remarks on truth and paradox. In P. A. French and H. K. Wettstein, editors, *Midwest Studies in Philosophy Volume XXXII: Truth and its Deformities*. Wiley-Blackwell, Boston, 2008.
- [Bra89] Ross T. Brady. The non-triviality of dialectical set theory. In G. Priest, R. Routley, and J. Norman, editors, *Paraconsistent Logic: Essays on the Inconsistent*, pages 437–470. Philosophia Verlag, Munich, 1989.
- [Bur79] Tyler Burge. Semantical paradox. *Journal of Philosophy*, 76:169–198, 1979. Reprinted in [Mar84].
- [Bur86] John P. Burgess. The truth is never simple. *Journal of Symbolic Logic*, 51:663–681, 1986.
- [BvF03] Jc Beall and Bas C. van Fraassen. *Possibilities and Paradox*. Oxford University Press, Oxford, 2003.
- [Can96] Andrea Cantini. *Logical Frameworks for Truth and Abstraction: An Axiomatic Study*. Elsevier, Amsterdam, 1996.
- [Chi79] Charles Chihara. The semantic paradoxes: A diagnostic investigation. *Philosophical Review*, 88:590–618, 1979.
- [Coo06] Roy Cook. There are non-circular paradoxes (but Yablo’s isn’t one of them). *The Monist*, 89:118–149, 2006.
- [Dow84] Bradley H. Dowden. Accepting inconsistencies from the paradoxes. *Journal of Philosophical Logic*, 13:125–130, 1984.
- [Dum59] Michael Dummett. Truth. *Proceedings of the Aristotelian Society*, 59:141–162, 1959. Reprinted in [Dum78].
- [Dum78] Michael Dummett. *Truth and Other Enigmas*. Harvard University Press, Cambridge, 1978.

- [Dun] J. Michael Dunn. Natural language versus formal language. Presented at the joint APA–ASL symposium, New York, December 27, 1969.
- [Ekl02] Matti Eklund. Inconsistent languages. *Philosophy and Phenomenological Research*, 64:251–275, 2002.
- [Fef84] Solomon Feferman. Toward useful type-free theories, I. *Journal of Symbolic Logic*, 49:75–111, 1984. Reprinted in [Mar84].
- [Fef91] Solomon Feferman. Reflecting on incompleteness. *Journal of Symbolic Logic*, 56:1–49, 1991.
- [Fie94] Hartry Field. Deflationist views of meaning and content. *Mind*, 103:249–285, 1994.
- [Fie08] Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.
- [FS87] Harvey Friedman and Michael Sheard. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33:1–21, 1987.
- [Gai88] Haim Gaifman. Operational pointer semantics: Solution to self-referential puzzles I. In Moshe Y. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 43–59. Morgan Kaufmann, Los Altos, 1988.
- [Gai92] Haim Gaifman. Pointers to truth. *Journal of Philosophy*, 89:223–261, 1992.
- [Gau06] Christopher Gauker. Against stepping back: A critique of contextualist approaches to the semantic paradoxes. *Journal of Philosophical Logic*, 35:393–422, 2006.
- [GB93] Anil Gupta and Nuel Belnap. *The Revision Theory of Truth*. MIT Press, Cambridge, 1993.
- [Gla01] Michael Glanzberg. The Liar in context. *Philosophical Studies*, 103:217–251, 2001.

- [Gla04a] Michael Glanzberg. A contextual-hierarchical approach to truth and the Liar paradox. *Journal of Philosophical Logic*, 33:27–88, 2004.
- [Gla04b] Michael Glanzberg. Quantification and realism. *Philosophy and Phenomenological Research*, 69:541–572, 2004.
- [Gla04c] Michael Glanzberg. Truth, reflection, and hierarchies. *Synthese*, 142:289–315, 2004.
- [Gla06] Michael Glanzberg. Context and unrestricted quantification. In A. Rayo and G. Uzquiano, editors, *Absolute Generality*, pages 45–74. Oxford University Press, Oxford, 2006.
- [Gri91] Patrick Grim. *The Incomplete Universe*. MIT Press, Cambridge, 1991.
- [Gro94] Willem Groeneveld. Dynamic semantics and circular propositions. *Journal of Philosophical Logic*, 23:267–306, 1994.
- [Gup82] Anil Gupta. Truth and paradox. *Journal of Philosophical Logic*, 11:1–60, 1982. Reprinted in [Mar84].
- [Hal97] Volker Halbach. Tarskian and Krippean truth. *Journal of Philosophical Logic*, 26:69–80, 1997.
- [Hec97] Richard Heck. Tarski, truth, and semantics. *Philosophical Review*, 106:533–554, 1997.
- [Her67] Hans G. Herzberger. The truth-conditional consistency of natural language. *Journal of Philosophy*, 64:29–35, 1967.
- [Her70] Hans G. Herzberger. Paradoxes of grounding in semantics. *Journal of Philosophy*, 67:146–167, 1970.
- [Her82] Hans G. Herzberger. Notes on naive semantics. *Journal of Philosophical Logic*, 11:61–102, 1982. Reprinted in [Mar84].
- [Hor90] Paul Horwich. *Truth*. Basil Blackwell, Oxford, 1990.
- [Koo92] Robert C. Koons. *Paradoxes of Belief and Strategic Rationality*. Cambridge University Press, Cambridge, 1992.

- [Kri75] Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975. Reprinted in [Mar84].
- [LS04] Greg Littmann and Keith Simmons. A critique of dialetheism. In Graham Priest, Jc Beall, and Bradley Armour-Garb, editors, *The Law of Non-Contradiction*, pages 314–335. Oxford University Press, 2004.
- [Mar84] Robert L. Martin, editor. *Recent Essays on Truth and the Liar Paradox*. Oxford University Press, Oxford, 1984.
- [McG91] Vann McGee. *Truth, Vagueness, and Paradox*. Hackett, Indianapolis, 1991.
- [MW75] Robert L. Martin and Peter W. Woodruff. On representing ‘true-in-L’ in L. *Philosophica*, 5:217–221, 1975. Reprinted in [Mar84].
- [Par74] Charles Parsons. The Liar paradox. *Journal of Philosophical Logic*, 3:381–412, 1974. Reprinted in [Par83].
- [Par83] Charles Parsons. *Mathematics in Philosophy*. Cornell University Press, Ithaca, 1983.
- [Par84] Terence Parsons. Assertion, denial, and the Liar paradox. *Journal of Philosophical Logic*, 13:137–152, 1984.
- [Pat07] Douglas Patterson. Understanding the Liar. In Jc Beall, editor, *The Revenge of the Liar*, pages 197–224. Oxford University Press, Oxford, 2007.
- [Patng] Douglas Patterson. Inconsistency theories of semantic paradox. *Philosophy and Phenomenological Research*, forthcoming.
- [Pri84] Graham Priest. Logic of paradox revisited. *Journal of Philosophical Logic*, 13:153–179, 1984.
- [Pri97] Graham Priest. Yablo’s paradox. *Analysis*, 57:236–242, 1997.
- [Pri06] Graham Priest. *In Contradiction*. Oxford University Press, Oxford, second edition, 2006.

- [Pri08] Graham Priest. *An Introduction to Non-Classical Logic*. Cambridge University Press, Cambridge, second edition, 2008.
- [Rea02] Stephen Read. The Liar paradox from John Buridan back to Thomas Bradwardine. *Vivarium*, 40:189–218, 2002.
- [Rea06] Stephen Read. Symmetry and paradox. *History and Philosophy of Logic*, 27:307–318, 2006.
- [Res08] Greg Restall. Modal models for Bradwardine’s theory of truth. *Review of Symbolic Logic*, 1:225–240, 2008.
- [RTG08] S. Rahman, T. Tulenheimo, and E. Genot, editors. *Unity, Truth and the Liar: The Modern Relevance of Medieval Solutions to the Liar Paradox*. Springer Verlag, Berlin, 2008.
- [RU06] Agustín Rayo and Gabriel Uzquiano, editors. *Absolute Generality*. Oxford University Press, Oxford, 2006.
- [Sha04] Stewart Shapiro. Simple truth, contradiction, and consistency. In Graham Priest, Jc Beall, and Bradly Armour-Garb, editors, *The Law of Non-Contradiction*. Oxford University Press, 2004.
- [Sha06] Lionel Shapiro. The rationale behind the revision theory. *Philosophical Studies*, 129:477–515, 2006.
- [Shang] Lionel Shapiro. Revenge and expressibility, forthcoming.
- [Sim93] Keith Simmons. *Universality and the Liar*. Cambridge University Press, Cambridge, 1993.
- [Soa99] Scott Soames. *Understanding Truth*. Oxford University Press, Oxford, 1999.
- [Sor98] Roy Sorensen. Yablo’s paradox and kindred infinite Liars. *Mind*, 107(425):137–155, 1998.
- [Sor03] Roy Sorensen. *A Brief History of Paradox*. Oxford University Press, Oxford, 2003.

- [Tar35] Alfred Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, 1:261–405, 1935. References are to the translation by J. H. Woodger as “The Concept of Truth in Formalized Languages” in [Tar83].
- [Tar44] Alfred Tarski. The semantic conception of truth. *Philosophy and Phenomenological Research*, 4:341–375, 1944.
- [Tar83] Alfred Tarski. *Logic, Semantics, Metamathematics*. Hackett, Indianapolis, second edition, 1983. Edited by J. Corcoran with translations by J. H. Woodger.
- [vF68] Bas C. van Fraassen. Presupposition, implication, and self-reference. *Journal of Philosophy*, 65:136–152, 1968.
- [vF70] Bas C. van Fraassen. Truth and paradoxical consequence. In Robert L. Martin, editor, *Paradox of the Liar*, pages 13–23. Ridgeview, Atascadero, 1970.
- [Vis84] Albert Visser. Four valued semantics and the Liar. *Journal of Philosophical Logic*, 13:181–212, 1984.
- [Woo84] Peter W. Woodruff. Paradox, truth, and logic part 1: Paradox and truth. *Journal of Philosophical Logic*, 13:213–232, 1984.
- [Yab93a] Stephen Yablo. Hop, skip, and jump: The agnostic conception of truth. *Philosophical Perspectives*, 7:371–396, 1993.
- [Yab93b] Stephen Yablo. Paradox without self-reference. *Analysis*, 53:251–252, 1993.