Synthese
https://doi.org/10.1007/s11229-022-04009-x

ORIGINAL RESEARCH

In defence of PKF

Ian Rumfitt[1]

Received: 4 August 2021 / Accepted: 14 December 2022

Abstract
I advance arguments in favour of PKF as an articulation of a central sense of the predicate 'true', and show how it illuminates the relationship between that sense and the 'external' notion of truth found in such claims as 'An utterance of the Liar Sentence does not say anything, and so is not true'.

Keywords Formal theories of truth · Partial logic · PKF · Kripke's theory of truth

1 Introduction

In their article 'Axiomatizing Kripke's Theory of Truth', Volker Halbach and Leon Horsten present a formal theory which they call 'PKF' (Halbach & Horsten, 2006). The name is an acronym. 'K' stands for Kripke, for all the theorems of PKF hold in any of the fixed points which Saul Kripke described in his 'Outline of a Theory of Truth' (Kripke, 1975).[1] 'F' stands for Feferman. In work which circulated widely in the 1980s and was eventually published as part of his 'Reflecting on Incompleteness' (Feferman, 1991), Solomon Feferman produced a strictly classical axiomatic theory, KF, all of whose theorems are true in any of the models obtained by 'closing off' one of Kripke's fixed points (for this notion see Sect. 8 below). As a theory of truth, though, KF faces problems, some of which are attributable to the strict classicality of its underlying logic. Hence the 'P', which stands for 'partial': the logic of PKF is Partial Logic, which differs from the classical system in making provision for well-formed formulae which may be 'undefined', i.e. which may lack either of the classical truth values.

PKF now seems to be little loved, even by its creators. In his book of 2011, *The Tarskian Turn*, Horsten gave it only a provisional recommendation. The rules of PKF, he wrote, 'give a *partial* articulation of the meaning of the concept of truth…PKF

[1] More precisely: in any of the fixed points under the Strong Kleene interpretation of the connectives and quantifiers. See §2 below.

✉ Ian Rumfitt
ian.rumfitt@all-souls.ox.ac.uk

[1] All Souls College, Oxford OX1 4AL, UK

Published online: 21 January 2023

only looks good until [a] better theory comes along. We should surely hold open the possibility that some future stronger inferential truth theory may determine the meaning of the concept of truth even further or may determine it in a slightly different way' (Horsten, 2011, p.147; emphasis in the original). In his book, *Axiomatic Theories of Truth*, which first appeared in the same year, Halbach could not bring himself to join even this chorus of faint praise.[2] 'There is', he wrote there, 'a substantial price to be paid for replacing classical logic with the non-classical logic of PKF' (Halbach, 2011, p.294 = Halbach, 2014, p.280)—a price which Halbach suggests he is no longer willing to pay.

Neglected by its parents as PKF may be, I wish to speak in its defence. For reasons to be set out, I think it articulates rather well one central sense of the word 'true'. Moreover, if the purpose of PKF is to articulate that sense, its use of a non-classical logic is appropriate (perhaps, indeed, inevitable) and does not carry a high price. To see its merits, though, we need to distinguish between different goals for which axiomatic theories of truth might be constructed. While PKF serves some of those goals poorly, it is well suited to other, equally legitimate purposes.

2 Semantic versus axiomatic theories of truth

Logical theories of truth, Feferman observed,

> are of roughly two kinds, semantical (or definitional) and axiomatic. Tarski (1935) inaugurated semantical theories with his definition of truth for a logically circumscribed language within a metalanguage for it, i.e. in a typed setting. He argued that this was necessary since a language which contains its own truth predicate is inconsistent if it satisfies a few basic assumptions, namely the T-scheme, classical propositional logic, and the capacity to form self-referential statements. However, the ordinary use of truth in natural language is untyped and the constraints of a hierarchical theory seem unduly restrictive…Thus it was that beginning in the 1960s, attempts were made to obtain useful consistent untyped semantical theories of truth by giving up some part of Tarski's basic assumptions. One of the most influential of these was due to Kripke (1975), who defined a notion of truth for an untyped three-valued language (Feferman, 2012, p.182).

Kripke's own exposition was informal, and it will help to apply his semantic theory to a specific formalized language. Let L be the language of arithmetic with the single atomic predicate '$\xi = \zeta$', with connectives '$\wedge$', '$\vee$', and '$\neg$', and unary quantifiers '$\forall$' and '$\exists$'. Let L_T be the language got by adding a one-place predicate '$T(\xi)$' to L. The terms of L include canonical numerals. We postulate a fixed system G of Gödel numbering, so that any well-formed expression e of L_T may be designated by the canonical numeral $\ulcorner e \urcorner$ of its Gödel number under G.

While Kripke assumes that any closed sentence in the base language L is bivalent, he allows that some involving '$T(\xi)$' may be 'undefined'. He treats '$T(\xi)$' as initially

[2] I shall give page references to both the original and revised versions of Halbach's book. The changes between the two concern technical matters and do not touch Halbach's assessment of PKF.

uninterpreted, and describes a construction which generates a variety of extension/anti-extension pairs $<S_1, S_2>$ as semantic values for it, each with the property that a closed sentence φ of L_T belongs to S_1 if and only if $T(\ulcorner\varphi\urcorner)$ belongs to S_1 and belongs to S_2 if and only if $T(\ulcorner\varphi\urcorner)$ belongs to S_2. Such a pair is called a *fixed point* for '$T(\xi)$'. Kripke took it to be essential, if '$T(\xi)$' is to be interpreted as 'ξ is true', that φ should be assessed as true (or false) if and only if $T(\ulcorner\varphi\urcorner)$ is so assessed. When '$T(\xi)$' has a fixed point as its semantic value, this condition will be met.

Because some closed sentences of L_T are undefined, the standard truth tables do not give a full account of the conditions in which they are true, or false: we also have to say how the presence of an undefined part affects the truth or falsity of the whole. Different accounts generate different fixed points and Kripke does not commit himself to a particular scheme. Like most commentators, I confine attention to the fixed points generated by the 'Strong Kleene' scheme (Kleene, 1952, p.332*f*).

Assuming that scheme, we can give a mathematically rigorous description of those fixed points. Given any disjoint pair of sets of sentences $<X_1, X_2>$ (= X) as a base, the *Kripke jump* operation, Φ_X, maps a pair of sets of sentences $<S_1, S_2>$ to a new such pair $<S_1{}^+, S_2{}^+>$ according to the following rule, in which sentences are identified with their Gödel numbers under G (compare Halbach, 2014, p.189, Definition 15.5):

$$
\begin{aligned}
S_1^+ \;=\; & X_1 \cup \\
& \{n :\ n \text{ is } \varphi \wedge \psi, \text{ where } \varphi \in S_1 \text{ and } \psi \in S_1\} \cup \\
& \{n :\ n \text{ is } \varphi \vee \psi, \text{ where } \varphi \in S_1 \text{ and } \psi \in S_1\} \cup \\
& \{n : n \text{ is } \neg\varphi, \text{ where } \varphi \in S_2\} \cup \\
& \{n : n \text{ is } \forall x\varphi, \text{ where } \varphi(t/x) \in S_1, \text{ for every constant term } t\} \cup \\
& \{n : n \text{ is } \exists x\varphi, \text{ where } \varphi(t/x) \in S_1, \text{ for every constant term } t\} \cup \\
& \{n : n \text{ is } T(\ulcorner\varphi\urcorner), \text{ where } \varphi \in S_1\}. \\
S_2^+ \;=\; & X_2 \cup \\
& \{n :\ n \text{ is } \varphi \vee \psi, \text{ where } \varphi \in S_2 \text{ and } \psi \in S_2\} \cup \\
& \{n :\ n \text{ is } \varphi \vee \psi, \text{ where } \varphi \in S_2 \text{ and } \psi \in S_2\} \cup \\
& \{n : n \text{ is } \neg\varphi, \text{ where } \varphi \in S_1\} \cup \\
& \{n : n \text{ is } \forall x\varphi, \text{ where } \varphi(t/x) \in S_2, \text{ for every constant term } t\} \cup \\
& \{n : n \text{ is } \exists x\varphi, \text{ where } \varphi(t/x) \in S_2, \text{ for every constant term } t\} \cup \\
& \{n : n \text{ is } T(\ulcorner\varphi\urcorner), \text{ where } \varphi \in S_2\}.
\end{aligned}
$$

Where the pair $< S_1, S_2>$ is disjoint, $< S_1{}^+, S_2{}^+>$ is disjoint too. X_1 (X_2) is the set of sentences assumed true (false) in the base theory.

Kripke's construction depends on the fact that Φ_X is *monotonic* for any base pair X: whenever $<S_1, S_2> \subseteq <T_1, T_2>$, $\Phi_X(<S_1, S_2>) \subseteq \Phi_X(<T_1, T_2>)$. It is this which ensures that the Φ_X operator has a variety of fixed points.[3] That is, there exist pairs of sets of sentences $<S_1, S_2>$ for which $\Phi_X(<S_1, S_2>) = <S_1, S_2>$. In any fixed point,

[3] Tarski's original fixed-point theorem (Tarski 1955) applied only to complete lattices. For generalizations to complete partial orders, see chapter 8 of Davey & Priestley 2002 (esp. p.188) and the references therein.

$\varphi \in S_1$ if and only if $T(\ulcorner\varphi\urcorner) \in S_1$ and $\varphi \in S_2$ if and only if $T(\ulcorner\varphi\urcorner) \in S_2$. That is, a fixed point of the jump operator has the property which Kripke took to be crucial if '$T(\xi)$' is to be interpreted as 'ξ is true' (Kripke, 1975, p.71).

3 The logic of Kripke's semantical theory of truth

What is the logic of Kripke's semantical theory of truth? The question seems to pose a conundrum. Because L_T has the resources to describe its own syntax, it will contain 'Liar' sentences which 'say of themselves' that they are not true. More precisely, there exists in L_T a sentence λ which is provably equivalent to $\neg T(\ulcorner\lambda\urcorner)$. Such a sentence is undefined in any fixed point for '$T(\xi)$', so Kripke's semantic theory of truth evaluates λ as undefined. Under the Strong Kleene rules, when λ is undefined, so is $\lambda \vee \neg\lambda$. So classical logic, in which every instance of the schema $\varphi \vee \neg\varphi$ must be true, appears to be precluded. On the other hand, Partial Logic seems quite unsuitable, too. Commenting on Dana Scott's axiomatization of that logic, Feferman noted the absence of the usual rule for introducing a conditional connective (the 'Deduction Theorem') and complained that in consequence '*nothing like sustained ordinary reasoning can be carried on in [partial] logic*' (Feferman, 1984, p.264; emphasis in the original). Certainly, Partial Logic is ill-suited for the sustained mathematical reasoning needed to prove Tarski's fixed-point theorem and its subsequent generalizations.

The resolution of the apparent tension is simple. The use of classical logic would be unjustified if a formula such as λ were being *used* in the semantical theory. As the formal exposition of Kripke's informal account makes clear, though, in his semantical theory the closed sentences of L_T are only *mentioned*. Each such sentence is assigned to one of three categories: true, false, or undefined. In formalizing the construction, there is no need to *use* the sentences of L_T; it suffices to use (canonical) *designators* for them. The fact, then, that some of those sentences are assessed as undefined does not threaten the use of classical logic within the semantical truth theory. That theory does not entail that every closed sentence is either true or false, but that is no more of a threat to classical *logic* (as opposed to classical semantics) than is the failure of a theory of colour to entail that every coloured object is either black or white. Because the logic of Kripke's semantical theory is classical, though, it does entail that every sentence is either true or not true.

In this, I am agreeing with a key claim of Kripke's. In *n*.18 of 'Outline', he records his amazement at hearing 'my use of the Kleene valuation compared occasionally to the proposals of those who favor abandoning standard logic "for quantum mechanics", or positing extra truth values beyond truth and falsity, etc.' (Kripke, 1975, p.64). When he writes that 'all our considerations can be'—and, he implies, ought to be—'formalized in a classical metalanguage' (*op. cit.*, p.65), he is right—assuming that the 'considerations' in question are those which pertain to his *semantical* theory of truth.

4 Reasons for axiomatizing a theory of truth

So much, for the moment, about that theory. We have next to ask why anyone might want to axiomatize it. The answer cannot be a desire for rigour. While Kripke's account of his semantical theory was informal, it can be made fully rigorous. So why axiomatize?

In the paper of 2012 from which I have already quoted, Feferman gives eight reasons for doing so. I start with the last two:

7. Given axiomatizations suggest natural variants such as by extending general principles from one's base theory (e.g. induction in arithmetic, or separation in set theory) to the theory with a truth predicate.
8. A given philosophical conception of truth may suggest a semantical construction or an axiomatization, and once made more explicit in the latter way, we are in a better position to assess the underlying conception (Feferman, 2012, p.183).

In nuce, my argument will be as follows. These two reasons for axiomatizing give rise to different desiderata on the resulting theories. While PKF may score poorly on those which emerge from the Seventh Reason, it does much better on desiderata pertaining to the Eighth.

A simple example illustrates the Seventh Reason. Let us suppose that the base theory is classical first-order Peano Arithmetic, *PA*. Let *Con*(*PA*) be the sentence in the language of arithmetic which canonically formalizes the statement that *PA* is consistent. By Gödel's Second Incompleteness Theorem, *PA* cannot prove *Con*(*PA*) (assuming that *PA* is consistent). Consider, though, the axiomatic truth theory *PA*+ which is got from *PA*, first, by adding the truth predicate '$T(\xi)$' to its language, and then by adding the axioms which comprise the 'positive inductive definition' of truth (for a statement of these, see McGee, 2005, p. 95). Assuming that mathematical induction is applicable to properties expressible in the expanded language, *PA*+ does prove *Con*(*PA*). In this way, the construction of an axiomatic truth theory brings metamathematical benefits: we can use it to justify mathematical and metamathematical claims which the base theory is unable to prove. 'Adjoining a truth predicate to the language of arithmetic and permitting it to appear within induction axioms permits us to prove the Gödel sentence, *Con*(*PA*), and other useful theorems besides' (McGee *op. cit.*, p.94).

Feferman's own motivation for constructing KF was a more ambitious and sophisticated application of this Reason. Gödel's Incompleteness Theorems imply that some of what we implicitly accept in accepting *PA*—the Gödel sentence for *PA*, for example, or *Con*(*PA*)—is not explicit in *PA* itself. To borrow a term from Robert Brandom, there is then the project of *making explicit* what we implicitly accept; and then of making explicit what is implicit in accepting the new explicit theory; and then of making explicit…. If this process reaches a fixed point, we arrive at what Feferman calls the *reflective closure* of the original theory. Before his 1991 paper, the standard way of capturing the reflective closure of an incomplete theory was via an ordinally indexed sequence of increasingly strong theories, and a corresponding sequence of predicates meaning 'true-in-such-and-such a theory', viz., the hierarchy RT_α of ramified truth predicates. The technical achievement in 'Reflecting on Incompleteness' was to present an axiomatic theory, KF, for a single truth predicate, within which ramified truth predicates indexed by ordinals up to ε_0 could be defined.

Feferman's Sixth Reason for axiomatizing truth theories was this:

6. One can compare like and unlike axiomatizations as to their proof-theoretical strength using an extensive body of well-established metamathematical techniques (Feferman, 2012, p.183).

It is here, I believe, that we find the key to Halbach's disillusion with PKF. Like Feferman's, Halbach's motivation for constructing an axiomatic truth theory was primarily metamathematical. However, when we make the comparison suggested in the Sixth Reason between PKF, which is a theory in Partial Logic, and its classical cousin KF, PKF comes off worse:

> [I]t is *here* where PKF reveals its shortcoming: the usual proof of transfinite induction up to ε_0 cannot be carried out in PKF. In Lemma 16.19 I proved transfinite induction up to any ordinal up to ω^ω, but it cannot be pushed any further…because of the absence in the non-classical system PKF of rules available in classical logic (Halbach, 2011, p.293 = Halbach, 2014, p.279; emphasis in the original).

Halbach, too, now takes the absence of certain classical logical rules to be a serious demerit of his and Horsten's creation.

A debate continues over whether Halbach was too pessimistic about the prospects for using PKF in executing interesting metamathematical projects; see Nicolai, 2018, Halbach & Nicolai, 2018, and Field, 2022. Some of this debate concerns the question whether Halbach was right to ascribe any limitations in this direction to PKF's underlying logic, Partial Logic, rather than to its specific axioms. There is also the complicating factor that Feferman himself came to take a quite different approach to the problem of characterizing the reflective closure of a schematic axiom system: in some later papers, he defined this using a notion of 'the unfolding of a schematic system' which made no use of any truth theory (see Feferman & Strahm, 2000, 2010). Interesting as these matters are, they are beside the present point. For my thesis is that, even if PKF is sub-optimal as a fulfilment of Feferman's Seventh Reason, it has a strong claim to meet the Eighth. That is, there is an attractive 'philosophical conception of truth' which PKF makes explicit, and 'once made more explicit…, we are in a better position to assess the underlying conception' (Feferman, 2012, p.183).

5 Strawsonian truth and contentless sentences

The conception I have in mind was put forward by P.F. Strawson as

> something uncontroversial and fairly general about truth. One who makes a statement or assertion makes a true statement if and only if things are as, in making that statement, he states them to be. Or again: one who expresses a supposition expresses a true supposition if and only if things are as, in expressing that supposition, he expressly supposes them to be (Strawson, 1970, p.180).

More briefly:

> A sentence is true if and only if things are as it says they are.[4]

There is a cognate account of falsity:

> A sentence is false if and only if things are not as it says they are.

Two observations about this conception of truth and falsity are relevant. First, when so glossed, 'Sentence *A* is true' is naturally heard as *presupposing* that there is a way *A* says things are, i.e. as presupposing that *A* expresses a complete proposition. Someone who asserts 'Things are as *A* says they are' takes it for granted, and expects the audience also to take it for granted, that *A* says something. The usual linguistic tests for presupposition confirm this. 'If things are as Mary's statement says they are, then John lied in the witness box' implies that there is a way Mary's statement says things are. So does 'Things are not as Mary's statement says they are'. It follows from the last point that '*A* is false' (when given Strawson's gloss) also presupposes that *A* says something.

In fact, a second observation is warranted. Stephen Yablo has usefully distinguished between 'catastrophic' and 'non-catastrophic' failures of presupposition (Yablo, 2006). A catastrophic failure deprives the statement which carries the presupposition of content, while a non-catastrophic failure does not. Yablo argues convincingly that many of the stock examples of presupposition failure are non-catastrophic. A speaker who says 'The present king of France is bald', for example, is likely to presuppose (i.e. to take for granted and expect his audience to take for granted) that France has a king. All the same, the fact that France has no king does not deprive his statement of a propositional content. To the contrary, we know full well how things would have to be for it to be true. Matters stand differently, however, with ascriptions of Strawsonian truth and falsity. Someone who asserts 'Things are as *A* says they are' presupposes that *A* says something. But if that presupposition fails—i.e. if *A* fails to express a proposition—then the statement 'Things are as *A* says they are' also fails to say anything. If there is no way *A* says things are, there is nothing left in the claim that things are that way. The parallel result obtains for falsity. If there is no way that *A* says things are, there is equally nothing left in the claim that things are not that way. Failures of the presuppositions carried by ascriptions of Strawsonian truth and falsity, then, are catastrophic in Yablo's sense.

When might the utterance of a meaningful declarative sentence fail to say that things are thus-and-so? Philosophers have suggested various examples. For present purposes, though, I focus on an argument which J.L. Mackie gave to show that certain versions of the Liar Paradox fail to express propositions. Consider the sentence 'No sentence uttered by a Cretan, standardly construed, makes a true statement', as uttered by Epimenides, a Cretan. Mackie writes:

[4] Strawson would have deprecated as barbarous ascriptions of truth to (declarative) sentences, as opposed to the statements made by uttering or inscribing such sentences. However, in the formal language L_T (cf. §2), any two inscriptions of the same closed sentence say the same thing, if they succeed in saying anything.

> Let us assume, what might be the case, that no other sentence uttered by a Cretan, standardly construed, makes a true statement. We cannot without contradicting ourselves allow that Epimenides's remark makes a true statement. And yet if it fails *for whatever reason* to make a true statement, we must ourselves say exactly what Epimenides has said; how then can we deny that this is a sentence uttered by a Cretan which, standardly construed, makes a true statement? How can we avoid contradicting ourselves? (Mackie, 1973, p.294).

> In addressing these questions, Mackie applies Strawson's explanation of truth:

> Suppose that we expand 'true' here, replacing 'would make a true statement' with 'would state that things are as they in fact are'. And remember that the things in question include the success or failure of this sentence itself in this respect. I think we can and must say that because of the very tricky kind of self-reference and consequent self-dependence in this case, there just is no *how things are* in the key respect. Consequently, we cannot either endorse or deny a sentence-token of the same type and with the same reference as [Epimenides's remark]…We must just admit that the issue it appears to raise is indeterminate, and hope our study of self-reference has explained why this is so. This sentence's indeterminacy with respect to truth is of a kind which prevents our saying even that it is not true, and therefore from arguing, by a further step, that it is true (Mackie, 1973, p.295).

I shall revert (in Sect. 8) to Mackie's final claim—that 'this sentence's indeterminacy with respect to truth is of a kind which prevents our saying even that it is not true'. Whatever the eventual judgement on that, though, his earlier argument—that 'there just is no *how things are* in the key respect'—is compelling. Since the assumption that Epimenides expressed a truth leads directly to a contradiction, and since it is given that no other Cretan has said anything true, we must conclude that no sentence uttered by a Cretan, when standardly construed, makes a true statement. Let us suppose, though, that Epimenides's utterance succeeds in strictly and literally saying something—viz. that p. Given what his words mean, the proposition that p must be equivalent to the proposition that no sentence uttered by a Cretan, when standardly construed, makes a true statement. But we have already concluded that no sentence uttered by a Cretan, when standardly construed, makes a true statement, so if Epimenides's utterance succeeds in saying something, then things are as it says they are, so it must be true, after all. The only way of avoiding becoming embroiled in contradiction ourselves is to deny that Epimenides's remark expresses a proposition. In elaborating Strawson's conceptions of truth and falsity, then, we must allow for the fact that some meaningful sentences involving 'true' and 'false', when uttered in certain contexts, fail to say anything.

6 Elaborating Strawson's conception of truth

Suppose we set out to make those conceptions more explicit by formulating axioms which relate truth and falsity (as so conceived) to other central logical notions, such as conjunction, disjunction, and negation. The fact that some utterances of meaningful sentences involving 'true' and 'false' fail to say anything implies that the underlying logic of the resulting theory cannot be classical. For that theory will contain sentences in which 'true' and 'false' are used, not merely mentioned. In classical logic, any instance of Excluded Middle is true, and it is hard to see how 'Either *A* or not *A*' can express a truth when *A* fails to say anything.

This poses a problem of method. In making Strawson's conception explicit, we shall need to engage in some deductive reasoning, but which logic sets the standards for valid deduction? My way of dealing with this problem will be to proceed informally, reasoning in English, but noting the logical principles which are required at each stage of the elaboration. Each principle needed, it will turn out, is a sound rule in Partial Logic, and by applying them to the Strawsonian conception we may deduce all of the axioms of PKF. It is in this sense that PKF—that is, the combination of Partial Logic with Feferman's truth axioms—constitutes a consequent elaboration of that conception. I do not claim that this combination provides the *only* elaboration of that conception: indeed, it is hard to imagine what an argument for this claim would look like. For all that, PKF will emerge as *one* way of unpacking Strawson's conception of truth.[5]

John Burgess (2014) usefully distinguishes 'positive' from 'negative', and 'compositional' from 'decompositional' principles concerning truth and falsity. Positive compositional principles say when truth may be ascribed to complex sentences, given assignments of truth and falsity to their parts. The positive compositional principle for conjunction is straightforward. Suppose that two sentences, *A* and *B*, are true in a given context of utterance. (Henceforth, I take the context as given and held fixed.) Then things are as *A* says they, and things are as *B* says they are. Suppose *A* says that *p*, and *B* says that *q*. Since things are as *A* says they are, *p*; since things are as *B* says they are, *q*. By the logical rule of Conjunction Introduction, it follows that *p* and *q*. But any conjunction of *A* and *B* says that *p* and *q*, so things are as that conjunction says they are. That is, any conjunction of *A* and *B* is true. From Conjunction Introduction, then, together Strawson's account of truth, we have derived the following compositional principle: if *A* is true and *B* is true, then any conjunction of *A* and *B* is true. This conditional claim is to be read as licensing the inference from antecedent to consequent.

Positive *de*compositional principles say what follows from ascriptions of truth to complex sentences. To find such a principle for conjunction, suppose that a conjunction of *A* with *B* is true. By Strawson's account of truth, things are as such a conjunction says they. Now a conjunction of *A* with *B* says that *p* and *q*, so we may infer that *p* and *q*. Given the logical rule of Conjunction Elimination, we may further deduce that *p*. That is, things are as *A* says they are. Using Strawson's account once more, we may conclude that *A* is true. One positive decompositional principle for conjunction, then, licenses the inference from the premiss that a conjunction of *A* and *B* is true to the conclusion that *A* is true. A parallel argument licenses the inference from the same premiss to the conclusion that *B* is true.

[5] There may well be others, e.g. by replacing Partial Logic with FDE.

'Negative' principles deal with falsity rather than truth, and here the issues are more subtle. It is a feature of entailment that, just as truth transmits forwards, falsity transmits backwards. If *A* entails *B*, and *B* turns out to be false, *A*'s alethic status is settled: it too is false. The logical rule of Conjunction Elimination, then, combines with Strawson's account of falsity to yield the following negative compositional principle for conjunction: if *A* is false, then a conjunction of *A* and *B* is false; and if *B* is false, such a conjunction is false.

It may be objected that the principle about the backward transmission of falsity has exceptions when meaningful sentences may fail to say anything. Suppose Epimenides says 'Snow is black and no sentence uttered by a Cretan, standardly construed, makes a true statement'. According to the argument of the previous paragraph, this conjunction is false. However, an attribution of Strawsonian falsity to a sentence presupposes that it says something, and what does this sentence say? Since its second conjunct says nothing, there appear to be only two possibilities. Either the failure of the second conjunct to say anything implies that the whole conjunction also fails to say anything, in which case it cannot be false. Or the whole sentence says merely that snow is black, in which case it is not really a conjunctive statement. (For this argument, see Soames, 1999, pp.194, 195.)

This objection is fallacious. It confuses what Dummett called a statement's free-standing 'content' with its 'ingredient sense'. We must distinguish, he wrote,

> between knowing the meaning of a statement in the sense of grasping the content of an assertion of it, and in the sense of knowing the contribution it makes to determining the content of a complex statement in which it is a constituent: let us refer to the former as simply knowing the content of the statement, and to the latter as knowing its ingredient sense (Dummett, 1981, pp.446, 447).

These notions are conceptually distinct and it is often vital to separate them (see e.g. Evans, 1979, pp.200–203, and Stanley, 1997). The objection of the previous paragraph is another case where the distinction matters, for the correct reply to it is this. Epimenides's remark 'No sentence uttered by a Cretan, standardly construed, makes a true statement' has a null free-standing content: an unembedded utterance of it says nothing. That, though, does not imply that it contributes nothing to the content of a complex sentence of which it is a part. We may consistently hold, then, that the conjunctive sentence uttered by Epimenides says that snow is black and no sentence uttered by a Cretan, standardly construed, makes a true statement. Since snow is not black, things are not as that sentence says they are. So the sentence is indeed false.

Suppose, finally, that we accept the following generalized version of the backward transmission principle: if some premisses entail a false conclusion, at least one of them is false. This combines with the rule of Conjunction Introduction to yield the following negative decomposition principle: if a conjunction of *A* and *B* is false, then either *A* is false or *B* is false. We have, then, a complete suite of semantic principles for conjunction. Positively: a conjunction is true if and only if each conjunct is true. Negatively, a conjunction is false if and only if one or other conjunct is false.

Parallel arguments yield dual compositional and decompositional principles for disjunction. Positively: a disjunction is true if and only if one or other disjunct is true. Negatively: a disjunction is false if and only if each disjunct is false. These principles in turn yield truth tables for conjunction and disjunction, in which the asterisk indicates lack of a truth-value:

Conj	T	*	F
T	T	*	F
*	*	*	F
F	F	F	F

Disj	T	*	F
T	T	T	T
*	T	*	*
F	T	*	F

The positive compositional principle for conjunction yields the T in the top-left corner of the *Conj* table, while the negative compositional principle yields the Fs in the bottom row and the right-hand column. The two decompositional principles for conjunction imply that no T or F can appear elsewhere in *Conj*, hence the asterisks. These displays resemble the 'Strong Kleene' tables of three-valued logic. In our tables, though, the asterisk signifies an absence of Strawsonian truth or falsity, not a third truth value.

There are analogous compositional and decompositional principles for the quantifiers. In most natural languages, quantifiers are binary: one argument place is filled with a count noun or a mass term; a second is filled by a verb or adjectival phrase. For all that, it is not wrong to analyse 'Everything is material' as a unary universal quantification of the predicable 'ξ is material', and 'Something smells' as a unary existential quantification of 'ξ smells'. Arguments analogous to those given above for conjunction yield the principles that a universal quantification of a predicable (with respect to a given domain) is true if and only if the predicable is true of every object in the domain, and is false if and only if the predicable is false of some object there. The principles for the unary existential quantification are dual. The notions of being *true of* and being *false of* may be explained in terms of truth and falsity. Thus, following Evans, we may say that a predicable is true of (false of) an object, x, if and only if there is an extension of the language to which the predicable belongs in which a hitherto unused name, β, designates x and the sentence which results from inserting β into the empty place of the predicable is true (false). As Evans notes, this requires the assumption that 'for every object there is an extension of the language which contains a name for that object' but not that 'there is an extension of the language which contains a name for every object' (Evans, 1977, p.84).

We wish our account of truth to be applicable to semantically closed languages, i.e. to languages which predicate truth of their own sentences. Again holding the context fixed, we can state semantic principles for such predications. For suppose A is true. Any predication of truth to A says that A is true. So things are as such a predication says they are, which is to say, the predication is itself true. We have, then, a positive compositional principle: if A is true, then any predication of truth to A is true. Conversely, suppose that things are as a predication of truth to A says they are. Since such a predication says that A is true, it follows that A is true, so we also have the corresponding decompositional principle: if a predication of truth to A is true, then A is true. Suppose next that A is false. Then things are not as A says they are, so A is not true. Any predication of truth to A, however, says that A is true, so things are not as such a predication says they are. That is, such a predication is false. This yields a negative compositional principle: if A is false, then any predication of truth to A is false. Finally, suppose that a predication of truth to A is false. Then things are not as such a predication says they are. Such a predication says that A is true, so it is not the case that A is true. That is, things are not as A says they are, i.e. A is false. So we have a negative decompositional principle: if a predication of truth to A is false, then A is false.

The argument in favour of this final principle may seem like sleight of hand. We are allowing for the possibility that a sentence may fail to say anything. Surely, it may be protested, such a sentence is neither true nor false, so that the inference from 'It is not the case that A is true' to 'A is false' must be fallacious. If we stick to Strawsonian truth and falsity, however, the inference is sound. 'A is true' and 'A is false' both carry the presupposition that A says something, and each is deprived of content if that presupposition fails. We can, then, pass back and forth between the truth of 'A is not true' and the truth of 'A is false'. In fact, we can pass back and forth between the falsity of these sentences, too. It is again important to recall that in saying e.g. 'A predication of truth to A is false if and only if A is false', all that being claimed is the soundness of the inferences back and forth. Truth is not being ascribed to any complex sentence.

The objector may still protest that there must be *some* senses of 'not', 'true', and 'false' for which the inference from 'It is not the case that A is true' to 'A is false' is fallacious. I agree, but they do not include the senses of 'true' and 'false' which Strawson captured, and which this section is devoted to unfolding.

Finally, what are the Strawsonian semantic principles for negation? If A says that p, then any negation of A says that not p. So if things are not as A says they are, they are as a negation of A says they are. This yields a positive compositional principle: if a sentence is false, any negation of it is true. Conversely, if a negation of A is true, things are as that negation says they. Since any negation of A says that not p, it follows that not p. But A says that p, so things are not as A says they are, whence A is false. We have, then, a positive decompositional principle: if a negation of a sentence is true, the original sentence is false.

In deriving these two positive principles, no substantial assumptions about the logic of negation are needed. Matters are different when it comes to the negative principles. The rule of Double Negation Introduction licenses the inference from the premiss that p to the conclusion that not not p. Suppose, then, that A is true. Since A says that p, we have that p. By Double Negation Introduction, it follows that not not p. Since any negation of A says that not p, this implies that things are not as such a negation says they are. We thus have a negative compositional principle: if a sentence is true, any negation of it is false. The rule of Double Negation Elimination licenses the inference from the premiss that not not p to the conclusion that p. Suppose that a negation of A is false. Since any negation of A says that not p, and since things are not as such a negation says they are, we have that not not p. By Double Negation Elimination, it follows that p. Since A says that p, this implies that things are as A says they are. This yields a negative decompositional principle: if a negation of a sentence is false, the original sentence is true. Together, these four theses imply that negation is a kind of logical switch which toggles between truth and falsity. When A fails to express a content, so does any negation of A, and conversely, so we have the table:

	Neg
T	F
*	*
F	T

The logical rules used in deriving the negative semantic principles for negation are seriously contested. Intuitionistic logicians deny the unrestricted applicability of Double Negation Elimination; inverse intuitionistic logicians deny that of Double Negation Introduction.

7 PKF vindicated

Where does this leave matters? Although the logical principles for negation are controversial, all the rules used in elaborating the Strawsonian conception of truth are sound in Partial Logic. What is more, that logic is well suited to express the principles which unpack that conception. Those principles license *inferences*—from, e.g., the premisses that A is true and that B is true to the conclusion that any conjunction of A and B is true. They are, then, naturally formalized as sequents. In Partial Logic, theories are usually treated as deductively closed sets of sequents $X \Rightarrow Y$, where X and Y are finite (perhaps empty) sets of sentences. In Blamey's formalization, the sequent $X \Rightarrow Y$ is said to be *correct* just when (a) some member of Y is true if all the members of X are true, and (b) some member of X is false if all the members of Y are false (Blamey, 2002, esp. pp.333 and 336*f.*). A correct sequent, then, transmits truth forwards along the arrow and falsity backwards, just as we wanted. (Condition (b) does not follow from (a) without the further premiss that each member of $X \cup Y$ is either true or false.) We use the notation '$X \Leftrightarrow Y$' to mean '$X \Rightarrow Y$ and $Y \Rightarrow X$'. Thus the *double-headed* sequent $X \Leftrightarrow Y$ is correct precisely when (a) some member of Y is true if all the members of X are true, (b) some member of X is false if all the members of Y are false, (c) some member of X is true if all the members of Y are true, and (d) some member of Y is false if all the members of X are false.

When we collect together, and formalize in this style, the various compositional and decompositional principles argued for in the previous section, and add further principles giving truth and falsity conditions for the atomic sentences of the base language L, we reach the following set of sequents, in which 'Var(ξ)', 'Term(ξ)', 'Sent(ξ)', and 'For(ξ, ζ)' are predicates in L_T (defined via Gödel numbering) equivalent to 'ξ is a variable', 'ξ is a closed term', 'ξ is a sentence', 'ξ is a formula with only the variable ζ free':

1a Term(s), Term(t), val(s) = val(t) $\Rightarrow T(\ulcorner s = t \urcorner)$
1b Term(s), Term(t), $T(\ulcorner s = t \urcorner) \Rightarrow$ val(s) = val(t)
2a Sent(φ), Sent (ψ), $T(\ulcorner \varphi \urcorner) \wedge T(\ulcorner \psi \urcorner) \Rightarrow T(\ulcorner \varphi \wedge \psi \urcorner)$
2b Sent(φ), Sent(ψ), $T(\ulcorner \varphi \wedge \psi \urcorner) \Rightarrow T(\ulcorner \varphi \urcorner) \wedge T(\ulcorner \psi \urcorner)$
3a Sent(φ), Sent(ψ), $T(\ulcorner \varphi \urcorner) \vee T(\ulcorner \psi \urcorner) \Rightarrow T(\ulcorner \varphi \vee \psi \urcorner)$
3b Sent(φ), Sent(ψ), $T(\ulcorner \varphi \vee \psi \urcorner) \Rightarrow T(\ulcorner \varphi \urcorner) \vee T(\ulcorner \psi \urcorner)$

4a Var(x), Form (φ, x), $\forall y T(\ulcorner\varphi(y/x)\urcorner) \Rightarrow T(\ulcorner\forall x\varphi\urcorner)$
4b Var(x), Form (φ, x), $T(\ulcorner\forall x\varphi\urcorner) \Rightarrow \forall y T(\ulcorner\varphi(y/x)\urcorner)$
5a Var(x), Form (φ, x), $\exists y T(\ulcorner\varphi(y/x)\urcorner) \Rightarrow T(\ulcorner\exists x\varphi\urcorner)$
5b Var(x), Form (φ, x), $T(\ulcorner\exists x\varphi\urcorner) \Rightarrow \exists y T(\ulcorner\varphi(y/x)\urcorner)$
6a Term(t), $T(t) \Rightarrow T(\ulcorner T(t)\urcorner)$
6b Term(t), $T(\ulcorner T(t)\urcorner) \Rightarrow T(t)$
7a Sent(φ), $\neg T(\ulcorner\varphi\urcorner) \Rightarrow T(\ulcorner\neg\varphi\urcorner)$
7b Sent(φ), $T(\ulcorner\neg\varphi\urcorner) \Rightarrow \neg T(\ulcorner\varphi\urcorner)$

In Axiom 1, val(x) is a function taking (numerals of Gödel numbers of) terms to the numbers which they designate. In Axioms 4 and 5, where φ is a formula whose only free variable is 'x', and y is a natural number, $\ulcorner\varphi(y/x)\urcorner$ is got by replacing each occurrence of 'x' in φ with the canonical numeral designating y. If, finally, we add an axiom saying that any truth is a sentence, viz.

8 $Tx \Rightarrow$ Sent(x),

we reach precisely the characteristic axioms of PKF; see Halbach & Horsten, 2006, pp.692–3.[6]

What about falsity conditions for sentences of L_T? PKF offers no independent account of these: it *defines* 'A is false' to mean 'The negation of A is true'. As an elaboration of Strawsonian conception, this is suboptimal. In Strawson's gloss on 'false', the word 'not' is used, not mentioned, so Strawsonian falsity ought to be applicable to sentences in a language which lacks any sign for negation. This, however, is not a serious demerit of PKF for current purposes. For the 'definition' (*sic*) of 'A is false' combines with PKF to yield formalizations of the various negative compositional and decompositional principles of the previous section.

In Partial Logic, the sequent $A \Leftrightarrow \neg A$ is not contradictory. Rather, its correctness shows that A is undefined. Axioms 6 and 7 imply that the Liar sentence λ is undefined. By definition, λ is equivalent to $\neg T(\ulcorner\lambda\urcorner)$. That is, $\lambda \Leftrightarrow \neg T(\ulcorner\lambda\urcorner)$. Partial Logic permits substitution of proven equivalents within any sequent. By substituting $\neg T(\ulcorner\lambda\urcorner)$ for λ within $T(\ulcorner\lambda\urcorner) \Leftrightarrow T(\ulcorner\lambda\urcorner)$, then, we reach $T(\ulcorner\lambda\urcorner) \Leftrightarrow T(\ulcorner\neg T(\ulcorner\lambda\urcorner)\urcorner)$. By Axiom 7, this yields $T(\ulcorner\lambda\urcorner) \Leftrightarrow \neg T(\ulcorner T(\ulcorner\lambda\urcorner)\urcorner)$ and hence $T(\ulcorner\lambda\urcorner) \Leftrightarrow \neg T(\ulcorner\lambda\urcorner)$ by Axiom 6.[7] Since $T(\ulcorner\lambda\urcorner) \Leftrightarrow \neg T(\ulcorner\lambda\urcorner)$ is an instance of $\varphi \Leftrightarrow \neg\varphi$, this already shows that $T(\ulcorner\lambda\urcorner)$ is undefined in truth value. In Partial Logic, $\lambda \Leftrightarrow \neg T(\ulcorner\lambda\urcorner)$ yields $\neg\lambda \Leftrightarrow T(\ulcorner\lambda\urcorner)$. Hence $\neg\lambda \Leftrightarrow \neg T(\ulcorner\lambda\urcorner) \Leftrightarrow \lambda$, completing the proof that λ is also undefined.

[6] Since Partial Logic includes the rule of contraposition, PKF 8 yields $\neg$Sent(x) $\Rightarrow \neg Tx$, which may seem doubtful. For $\neg T(\ulcorner\varphi\urcorner)$, like $T(\ulcorner\varphi\urcorner)$, carries the presupposition that φ succeeds in saying something; and surely that presupposition fails when x is not even a declarative sentence. I claim, *per contra*, that PKF 8 is acceptable. When x is not a sentence—when for example, it is a tea-cup—there is no presupposition that things are as x says they are. 'That tea-cup is true' is straightforwardly false and 'That tea-cup is not true' is straightforwardly true. An ascription of truth to x only carries the presupposition that x says something when x is the sort of thing which *could* say something, and in the domain of PKF the only such things are closed sentences. (Thanks to a referee for pressing for clarification here.).

[7] Recall that '$X \Leftrightarrow Y$' means '$X \Rightarrow Y$ and $Y \Rightarrow X$'. When I write that a double-headed sequent follows 'by Axiom 7' (say), I mean that one component sequent follows by 7a and its converse follows by 7b.

8 Strawsonian truth versus external truth

This completes the positive case for PKF as an articulation of Strawson's conception of truth. Its only demerit is the lack of a parallel axiomatization of the cognate notion of falsity. I conclude by arguing that, as per Feferman's Eighth Reason, the axiomatization puts us 'in a better position to assess the underlying conception'. It does so by clarifying the relationship between Strawsonian truth and a cognate but distinct notion.

To see what that notion might be, let us revert to Mackie's version of the Liar Paradox (in Sect. 5 above). Mackie argues convincingly that Epimenides's utterance fails to say anything. He further contends that this failure—or as he puts it 'this sentence's indeterminacy with respect to truth'—'is of a kind which prevents our saying even that it is not true, and therefore from arguing, by a further step, that it is true' (*op. cit.*, p.295).

Although Mackie calls this 'awkward' (*ibid.*), I argued in Sect. 6 that it is the correct conclusion to draw *if* we understand 'true' in Strawson's way. That, though, cannot be the end of the matter. Epimenides's remark, Mackie asserts, fails to say anything. That is, it fails to satisfy a presupposition of truth. But in that case there has to be *some* sense in which the remark is not true. What is that sense?

Some may be tempted to articulate it by positing an ambiguity in 'not'. In Sect. 6 we identified a notion of negation, *Neg*, with the following truth table:

	Neg
T	F
*	*
F	T

If 'not' signifies *Neg*, then in 'Epimenides's remark is not true' we fail to say anything. Yet alongside this 'internal' negation some have discerned another, equally legitimate 'external' notion, *NEG*, whose truth table is:

	NEG
T	F
*	T
F	T

The English particle 'not', it may further be claimed, is ambiguous between the two notions. So long as we mean *NEG* by 'not', we can truly say 'Epimenides's remark is not true'.

The fundamental difficulty with this view is the absence of uncontentious cases where 'not' signifies *NEG* (see Tappenden, 1999). In their absence, the claim of ambiguity looks like a philosopher's fancy rather than anything grounded in the linguistic facts. We do better, I contend, to recognize an ambiguity in 'true'. Senses are not to be multiplied lightly, but other writers have suggested that 'true' has different meanings, discernible only after theoretical investigation, somewhat as the term 'mass' can mean 'rest mass' or 'inertial mass'.[8] On the view to be presented, moreover, the two senses of 'true' are far from being a case of mere polysemy. Strawson captured central or 'focal' senses of the words 'true' and 'false' as they apply to declarative sentences. Reflection on the theory which makes those senses explicit, though, generates secondary meanings for the words. It is in this secondary sense that one may correctly assert 'Epimenides's remark is not true'.

What relation does the secondary sense of 'true' bear to the one which Strawson captured and which PKF unpacks? As Fischer et al (2015) have noted, there is a sense in which PKF is not merely sound with respect to Kripke's semantical account of truth, but also constitutes a complete axiomatization of that theory.[9] For let **N** be the standard natural number structure, and take S to be an arbitrary set of natural numbers (which generates a corresponding set of sentences, viz., those whose Gödel numbers under G belong to S). If we interpret S to be the extension of the truth predicate '$T(\xi)$', it may be shown that for all $S \subseteq \omega$

$$(\mathbf{N}, S) \models_{\mathrm{PL}} \text{PKF if and only if } (\mathbf{N}, S) \in SK$$

where SK is the class of fixed points of Kripke's construction under the Strong Kleene scheme, and $\models_{\mathrm{PL}}$ signifies truth-in-a-model of Partial Logic (Fischer et al., 2015, p.274, Theorem 4.12). In this way, PKF may be said to determine the class of fixed points under the Strong Kleene scheme (for a given formalized language).

With that class of fixed points determined, we can proceed to determine the possible extensions of a secondary sense of the word 'true':

> Take any fixed point $L'(S_1, S_2)$. Modify the interpretation of $T(x)$ so as to make it false of any sentence outside S_1. [We call this 'closing off' $T(x)$.] A modified version of Tarski's Convention T holds in the sense of the conditional $T(\mathrm{k}) \vee T(\mathrm{neg(k)}) \,.\supset.\, A \equiv T(\mathrm{k})$. In particular, if A is a paradoxical sentence we can now assert $\neg T(\mathrm{k})$ (Kripke, 1975, pp.80, 81).

It is in this sense that Epimenides's sentence, λ, may correctly be said to be not true. In no fixed point $L'(S_1, S_2)$ does λ belong to S_1. So, when we 'modify the interpretation of ["true'] so as to make it false of any sentence outside S_1', we can rightly say that λ is not true.

This explanation of the secondary sense of 'true' involves heavy use of metalogic. We identify the class of models or fixed points delineated by PKF. We then apply a metalogical operation, that of 'closing off' those fixed points. This account of how the extension is determined gives us what grip we have on the secondary sense of 'true'. For this reason, we may follow Burgess (2014) in labelling this secondary sense 'external'. Burgess, in fact, 'tentatively' accepts PKF (which he labels 'KHH') as an axiomatization of the 'internal' sense of 'true' (2014, p.138). His informal gloss

[8] For this comparison, see Field 1994 and McGee 2005, although neither of these authors is concerned precisely with the difference between Strawsonian truth and what I shall call 'external' truth.

[9] There is, of course, no prospect of a complete axiomatization of the arithmetic which underpins the formal syntax. That is why I take the standard natural number structure, **N**, as given (cf. Halbach 2011, p.211 = Halbach 2014, p.197). We are concerned to axiomatize the properly semantical parts of Kripke's theory of truth, not the presuppositions of its syntax.

on that sense, though, and his argument for PKF as an unfolding of it, are very different from mine. (He does not mention Strawson.)

Can, though, the 'external' sense also be captured in axioms which unfold a philosophically attractive conception of truth? Feferman's original classical theory KF (Feferman, 1991) is sound and complete with respect to the class of closed-off models in precisely the sense in which PKF is sound and complete with respect to Kripke's partial models. That is, where *CSK* is the class of all closed-off fixed points under the Strong Kleene scheme, for all $S \subseteq \omega$,

$$(\mathbf{N}, S) \models_{\text{CL}} \text{KF if and only if } (\mathbf{N}, S) \in CSK$$

(Fischer et al., 2015, p.268, Theorem 4.2).

Because the underlying logic of KF is classical, $\lambda \vee \neg\lambda$ is a theorem. However, KF (with Consistency[10]) also proves $\neg T(\ulcorner \lambda \vee \neg\lambda \urcorner)$. Thus, 'in the Kripke-Feferman theory, we can prove things that are, according to the Kripke-Feferman theory, untrue' (McGee, 1991, p.106). McGee takes this severing of the expected connection between proof and truth to be a decisive objection to the claim that KF unfolds any natural notion of truth: 'we do not see why we prove things, if proving something gives us no reason to suppose it is true' (*ibid.*).

How serious is this problem? Halbach and Horsten give reasons against evading it by retreating to a version of KF without Consistency (Halbach & Horsten, 2006, Sect.2). They also (*op. cit.*, Sect.1) identify problems for Reinhardt's (1986) project of making an 'instrumentalist' use of KF (although see now Castaldo & Stern, 2022 for an attempt to revive that project). Tim Maudlin, though, tackles the objection head on (Maudlin, 2004, pp.97–101). On Kripke's account, $\lambda \vee \neg\lambda$ is ungrounded and hence not true in the external sense. According to Maudlin, however, an assertion of it may still be *permissible*. So, the fact that KF legitimates such an assertion by proving the sentence is not an objection to that theory.

A debate is called for, but it is not obvious that Maudlin disposes of McGee's objection. As Maudlin says, there is a conceptual distinction between asserting a sentence and ascribing truth to it. Since Frege, though, philosophers have often explained asserting *A* as presenting *A* *as* true. On this conception of assertion, it is unclear how asserting an ungrounded sentence can be permissible. McGee's problem returns, for KF presents (for example) $\lambda \vee \neg\lambda$ *as* true, while at the same time saying that it is not. While not formally contradictory, this is perplexing. Perhaps Maudlin has a different conception of asserting in mind, but until that is articulated and compared with the Fregean conception, the sense in which the assertion of an ungrounded sentence is permissible is left somewhat mysterious.

In KF, 'is false' is defined as 'has a true negation'. One might instead try to axiomatize external truth as part of a broader theory in which 'false' or 'undefined', or both, are also treated as primitive notions. This is the approach Feferman took in his later theory DT, which contains separate 'axioms for determinateness and truth' (Feferman, 2008). DT's underlying logic is classical, and it yields *T*-theorems in precisely

---

[10] I.e. KF with the axiom $\forall x(\text{Sent}(x) \supset \neg\, (T(x) \wedge T(\text{neg}(x)))$. Halbach & Horsten (2006, p.682) take Consistency to be a constituent axiom of KF, but more recent writers tend not to.

the form of Kripke's 'modified version of Tarski's Convention T' (see above). But it does not give us what we are looking for. One of DT's determinateness axioms says that a disjunction is determinate only if both disjuncts are. This matches the Weak Kleene scheme, not the Strong.[11]

John Burgess has proposed another classical theory, KFμ. Its axioms are the positive and negative compositional principles of KF, but in place of the decompositional principles Burgess has an axiom schema of minimality: for an arbitrary condition on sentences $\varphi(x)$, it is an axiom that if the set of truths satisfying $\varphi(x)$ is closed under the compositional principles, then every true sentence satisfies $\varphi(x)$.[12] However, although KFμ omits some axioms of KF, it contains all its theorems, so it too proves both $\lambda \vee \neg\lambda$ and $\neg T(\ulcorner \lambda \vee \neg\lambda \urcorner)$. Like KF, then, it breaks the conceptual connection between proving something and proving it true, a connection which, pending a new, non-Fregean account of asserting, any notion of truth must respect.

In the quest, then, to find axioms for the external sense of 'true' which unfold a philosophically coherent conception of the notion, the auguries are uncertain. Perhaps we can apprehend that sense, as it applies to sentences in a given language, only by learning to speak an essentially richer metalanguage. However that may be, the unfolding of Strawson's conception in the axioms of PKF gives us a clear picture of how that internal notion relates to its external cousin, and thereby meets Feferman's Eighth Reason for axiomatizing truth.

9 Prospects and problems

Axioms 4 and 5 of PKF unpack Strawsonian truth as it applies to formulae whose main operators are the unary quantifiers '$\forall$' and '$\exists$'. However, the English quantifiers 'every' and 'some' (and their correlates in other natural languages) are binary. Bruno Whittle (2019) has shown how Kripke's semantical theory of truth can be extended to cover binary 'every' and 'some' in a way which preserves monotonicity. In 'Generalized Quantification in an Axiomatic Truth Theory' (Rumfitt, forthcoming), I first expand Partial Logic to cover binary universal and existential quantifiers and show that the resulting system is sound and complete with respect to the natural extension of Blamey's semantics. I then extend PKF to the expanded language and show that the resulting theory, PKF+, determines Whittle's fixed point models in precisely the same sense in which PKF determines Kripke's.

This gives hope for a satisfactory treatment of the natural-language conditional in PKF. As Feferman noted (see again the quote from his 1984 paper in §3), the absence of a Deduction Theorem in Partial Logic precludes a conditional *connective* with the expected logical properties. The dominant contemporary view among empirical linguists, however, is that the English word 'if' and its translations in other languages are not connectives (see e.g. Kratzer, 1981). Rather, an 'if'-clause serves to restrict

[11] More precisely, DT combines Weak Kleene truth conditions for disjunction with stronger truth conditions for a conditional which is not equivalent to the material one. See Feferman (2008, §2).

[12] Burgess 2014, p. 140. Burgess writes as though the only standard model of KFμ is the closed-off minimal fixed point under the Strong Kleene scheme. This is wrong: Fischer et al.show that no classical axiomatization uniquely singles out the minimal fixed point: see Fischer et al. (2015, p. 268).

an explicit or implicit binary quantifier. On this view, the paradigm use of 'if' is that illustrated in David Lewis's 'adverbs of quantification' (Lewis, 1975): 'Mostly, if John attends the discussion group, there is a bust-up' has the semantic structure 'At most meetings of the discussion group which John attends, there is a bust-up'. The bare conditional 'If *A* then *B*' is taken to have the binary structure 'Every nearby *A*-situation is a *B*-situation'. By applying the treatment of binary quantifiers mentioned in the previous paragraph to these structures, we reach a simple account of the truth and falsity conditions of conditionals within PKF+. While some features of the account will offend some philosophers, the treatment is at least simpler than Hartry Field's efforts to extend PKF to the conditional (e.g. in Field, 2016), which are highly complex and consequently implausible as semantic principles which ordinary speakers could follow, even implicitly.

Further discussion of this and other extensions of PKF must await another paper. I hope to have shown, though, that PKF unfolds Strawson's conception of truth and thereby fulfils Feferman's Eighth Reason for axiomatizing truth. It should not be cast aside.[13]

Declarations

Conflict of interest There are no conflicts of interest pertaining to this submission.

References

Blamey, S. R. (2002). Partial logic. In D. M. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic* (2nd ed., Vol. 5, pp. 261–353). Kluwer.

Burgess, J. P. (2014). Friedman and the axiomatization of Kripke's theory of Truth. In N. Tennant (Ed.), *Foundational adventures: Essays in honour of Harvey M. Friedman* (pp. 125–148). London: College Publications.

Castaldo, L., & and Stern. J. (2022). KF, PKF, and Reinhardt's program. *The Review of Symbolic Logic*.

Davey, B. A., & Priestley, H. A. (2002). *Introduction to lattices and order* (2nd ed.). Cambridge University Press.

Dummett, M. A. E. (1981). *Frege: Philosophy of language* (2nd ed.). Duckworth.

Evans, M. G. J. (1977). Pronouns, quantifiers, and relative clauses (I). *The Canadian Journal of Philosophy, 7*, 476–536. Page references are to the reprint in Evans 1985, pp.76–152.

Evans, M. G. J. (1979). Reference and contingency. *The Monist, 62*, 161–89. Page references are to the reprint in Evans 1985, pp.178–213.

Evans, M. G. J. (1985). *Collected papers*. Clarendon Press.

Feferman, S. (1984). Toward useful type-free theories I. *The Journal of Symbolic Logic, 49*, 75–111. Page references are to the reprint in Martin, ed. 1984, pp.53–81.

[13] Thanks to Volker Halbach, Daniel Waxman and to two referees for their comments on a draft.

Feferman, S. (1991). Reflecting on incompleteness. *The Journal of Symbolic Logic, 56*, 1–49.
Feferman, S. (2008). Axioms for determinateness and truth. *The Review of Symbolic Logic, 1*, 204–217.
Feferman, S. (2012). Axiomatizing truth: Why and how? In U. Berger et al., eds., *Logic, Construction, Computation* (Frankfurt: Ontos), pp.181–96.
Feferman, S., & Strahm, T. (2000). The unfolding of non-finitist arithmetic. *Annals of Pure and Applied Logic, 104*, 75–96.
Feferman, S., & Strahm, T. (2010). The unfolding of finitist arithmetic. *The Review of Symbolic Logic, 3*, 665–689.
Field, H. H. (1994). Deflationist views of meaning and content. *Mind, 103*, 249–285.
Field, H. H. (2016). Indicative conditionals, restricted quantification, and naïve truth. *The Review of Symbolic Logic, 9*, 181–208.
Field, H. H. (2022). The power of naïve truth. *The Review of Symbolic Logic*.
Fischer, M., Halbach, V., Kriener, J., & Stern, J. (2015). Axiomatizing semantic theories of truth. *The Review of Symbolic Logic, 8*, 257–278.
Halbach, V. (2011). *Axiomatic theories of truth*. Cambridge University Press.
Halbach, V. (2014). *Axiomatic theories of truth* (revised edition). Cambridge University Press.
Halbach, V., & Horsten, L. F. (2006). Axiomatizing Kripke's theory of truth. *The Journal of Symbolic Logic, 71*, 677–712.
Halbach, V., & Nicolai, C. (2018). On the costs of nonclassical logic. *Journal of Philosophical Logic, 47*, 227–257.
Horsten, L. F. (2011). *The Tarskian turn: Deflationism and axiomatic truth*. MIT Press.
Kleene, S. C. (1952). *Introduction to metamathematics*. Van Nostrand.
Kratzer, A. (1981). The notional category of modality. In H.-J. Eikmeyer & H. Reiser (Eds.), *Words, worlds, and contexts* (pp. 38–74). Walter de Gruyter.
Kripke, S. A. (1975). Outline of a theory of truth. *The Journal of Philosophy, 72*, 690–716. Page References are to the reprint in Martin, ed., 1984, 53–81.
Lewis, D. K. (1975). Adverbs of quantification. In E. L. Keenan (Ed.), *Formal semantics of natural language* (pp. 3–15). Cambridge University Press.
McGee, V. (1991). *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*. Indianapolis: Hackett.
McGee, V. (2005). Two conceptions of truth?—Comment. *Philosophical Studies, 124*, 71–104.
Mackie, J. L. (1973). *Truth, probability, and paradox: Studies in philosophical logic*. Clarendon Press.
Martin, R. L. (Ed.). (1984). *Recent essays on truth and the liar paradox*. Clarendon Press.
Maudlin, T. (2004). *Truth and paradox: Solving the riddles*. Oxford University Press.
Nicolai, C. (2018). Provably true sentences across axiomatizations of Kripke's theory of truth. *Studia Logica, 106*, 101–130.
Reinhardt, W. N. (1986). Some remarks on extending and interpreting theories with a partial predicate for truth. *Journal of Philosophical Logic, 15*, 219–251.
Rumfitt, I. (Forthcoming). Generalized quantification in an axiomatic truth theory.
Soames, S. (1999). *Understanding truth*. Oxford University Press.
Stanley, J. (1997). Rigidity and content. In R. G. Heck, Jnr (Ed.), *Language, thought, and logic: Essays in honour of Michael Dummett* (pp. 131–156). Oxford: Clarendon Press.
Strawson, P. F. (1970). *Meaning and Truth. An Inaugural Lecture delivered before the University of Oxford on 5 November 1969*. Oxford: Clarendon Press. Page references are to the reprint in Strawson, *Logico-Linguistic Papers* (pp. 170–189). London: Methuen, 1971.
Tappenden, J. (1999). Negation, denial, and language change in philosophical logic. In D. M. Gabbay & H. Wansing (Eds.), *What is negation?* (pp. 261–298). Kluwer.
Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica* **I**: 261–405. Translated by J.H. Woodger as 'The concept of truth in formalized languages' in Tarski, eds. Woodger and J. Corcoran, *Logic, Semantics, Metamathematics* (pp. 152–278). Indianapolis: Hackett, 1983.
Tarski, A. (1955). A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics, 5*, 285–309.
Whittle, B. (2019). Truth and generalized quantification. *Australasian Journal of Philosophy, 97*, 340–353.
Yablo, S. J. (2006). Non-catastrophic presupposition failure. In J. J. Thomson & A. Byrne (Eds.), *Content and modality: Themes from the philosophy of Robert Stalnaker* (pp. 164–190). Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.