



Yannis Stephanou

CLASSICAL LOGIC AND THE LIAR

Abstract. The liar and kindred paradoxes show that we can derive contradictions when we reason in accordance with classical logic from the schema (T) about truth: **S** is true iff p , where ‘ p ’ is to be replaced with a sentence and ‘**S**’ with a name of that sentence. The paper presents two arguments to the effect that the blame lies not with (T) but with classical logic. The arguments derive contradictions using classical logic, but instead of appealing to (T), they invoke semantic claims that seem even harder to reject. The first argument relies on two standard semantic principles that are not disquotational and on the claim that if there is such a thing as the property of being true, then ‘true’ expresses that property. The second argument relies on a schema about meaning: **S** means that p , where ‘**S**’ and ‘ p ’ are to be replaced as before.

Keywords: liar paradox; truth; nonclassical logics

Introduction

The liar is the paradox of the sentence that denies its own truth. In a typical form, it concerns the sentence

(L) is not true. (L)

As is well-known, we can easily end up in contradiction about that sentence.

The contradiction arises when we apply the schema

S is true iff p . (T)

To get an instance of the schema, we must replace the letter ‘ p ’ with a declarative sentence that is not ambiguous or indexical and the letter ‘**S**’

with a name of that sentence. Ambiguity and indexicality induce complications. The name replacing ‘S’ may either consist in the declarative sentence itself put inside quotation marks or be different. However, I will call (T) ‘disquotational’ in the sense that, in each instance, an expression (sentence) is mentioned on the one side and used on the other. The schema (T) appears to be a principle that characterizes the concept of truth.

It seems that one instance of (T) is the biconditional

$$(\mathbf{L}) \text{ is true iff } (\mathbf{L}) \text{ is not true.}$$

Assume that (L) is true. Then, because of the biconditional, it is not true. Hence, by *reductio ad absurdum*, we can deny the assumption: (L) is not true. Consequently, because of the biconditional again, it is true — contradiction.

The paradox has many versions. It may involve circular reference instead of self-reference. And it cannot be confined to a specifiable set of odd sentences, such as (L). Ordinary sentences about truth that in normal circumstances involve no contradiction can, in unusual but possible circumstances, become as paradoxical as (L).

The liar and kindred paradoxes are the obstacle to developing a coherent and plausible formal theory of truth. One approach to the problem, developed in recent decades, is to impose no restriction on schema (T), but deviate instead from classical logic without accepting contradictions ‘*p* and not-*p*’. The idea is that we could endorse the biconditional ‘(L) is true iff (L) is not true’, but disable the inference from it to ‘(L) is true and (L) is not true’. The approach has mainly been expounded by H. Field in a series of papers culminating in his book [5]. Field has argued for it by comparing it with many other treatments of the liar. Others have constructed nonclassical logics that fall within the same approach but differ from Field’s.¹ This paper does not aim to adjudicate between different versions of the approach, but to support the general idea. I will present two arguments to the effect that the blame for the contradiction lies not with (T) but with classical logic.

It is not the most usual treatment of the problem to keep (T) in its unrestricted form at the expense of classical logic. It is more usual to impose some restriction or other on (T) (or on the corresponding schema about propositions or on the one about tokens of sentences). The effect

¹ See e.g. S. Yablo’s [27], E. Zardini’s [28] and my [20].

of the restriction is to reject instances of the schema that involve paradoxical sentences (or paradoxical propositions or tokens of sentences).² Those who limit (T) may or may not also deviate from classical logic.³ G. Priest's work is also an example of the relatively unusual treatment. In [13] and elsewhere, he endorses (T) unreservedly and goes against classical logic in an extreme way, since he accepts contradictions ' p and not- p '. The approach exemplified by Field does not go so far.

It may be tempting to consider that paradoxical sentences lack truth-value. Yet there is a serious difficulty with the view that (L) lacks truth-value. If it has no truth-value, (L) is not true (and, of course, it is not false either). I have just written (L) itself. The view that (L) has no truth-value commits us to asserting (L). Now, asserting (L) is incoherent. The problem is not that if one asserts '(L) is not true', one takes the first step towards a contradiction because in a short while one will be compelled to add '(L) is true'. Such a step is taken by whoever is willing to apply the schema (T) to (L). For (T) allows us to infer from '(L) is not true' to '(L) is true'. But one may believe that (L) has no truth-value, so it is not true, and also believe that (T) should not be applied to sentences having no truth-value and that we should not infer from such a sentence to calling it 'true'. Then, one avoids contradiction. The problem is, that at least provided a sentence is not ambiguous or indexical, it is incoherent to assert it and also assert that it is not true.

The incoherence is not a contradiction; one does not assert both the sentence and its negation, nor does one assert that the sentence is both true and not true. But it is similar to a contradiction. When some

² The analogue of (T) for propositions is the schema

$$\Pi \text{ is true iff } p,$$

where ' p ' is to be replaced with a sentence expressing a proposition and ' Π ' with a name of that proposition. The analogue of (T) for tokens of sentences is the schema

$$\mathbf{U} \text{ is true iff } p,$$

where its instances are not biconditional sentences of that form, but tokens of such sentences, and where, in each instance, ' p ' is replaced with a token \mathbf{t} of a declarative sentence and ' \mathbf{U} ' is replaced with a token of a name of \mathbf{t} . A. Tarski, in his classic [22], discusses languages that possess no paradoxical sentence and limits (T) to such languages. J. Barwise and J. Etchemendy, in [2], restrict the schema about propositions, at least when they discuss what they describe as the Russellian view of propositions. H. Gaifman, in [6], effectively restricts the schema about tokens of sentences.

³ For instance, B. Van Fraassen deviates in [23] despite his nominal adherence to classical logic. He rejects the rule that allows us to conclude $\lceil \neg \mathbf{A} \vee \mathbf{B} \rceil$ once we have inferred \mathbf{B} from \mathbf{A} . For he considers that $\lceil \neg \mathbf{A} \vee \mathbf{B} \rceil$ may be false while the inference of \mathbf{B} from \mathbf{A} is valid.

philosophers assert ‘(L) is not true’, they do both: they both assert a sentence – (L) – and assert that the sentence is not true. Of course, they do them by means of the same act, uttering (L), but at any rate they do them. The incoherence stems from the fact that truth is a basic norm governing assertion. If what is asserted is true, this counts as a success for the speaker: she is right. If it is false, this counts as a failure: she is wrong. Given how central truth is to our practice of evaluating assertions, whoever makes an assertion lays claim to saying something true. So it is incoherent to add that it is not true. The incoherence is independent of (T) and of any willingness to infer from a sentence to calling it ‘true’.⁴

In order to tackle the liar paradox, we need to deviate either from schemas like (T), by refusing to apply them to paradoxical sentences, or from classical logic, by refusing to follow all its rules when our reasoning contains such sentences. If we deviate from classical logic, we will continue to accept some of its rules even when reasoning with a paradoxical sentence, but we will not accept all of them. Of course, anyone who considers that diverging from classical logic must always be the last resort will tackle the dilemma by targeting schemas such as (T). But it seems to me too rigid to believe that when one faces a theoretical difficulty, one should deviate from classical logic only if there is absolutely no other way to overcome the difficulty.

The rules and principles that make up classical logic possess obviousness, to a greater or lesser extent. But the same holds good in the case of extralogical principles, such as schema (T) about truth. Let’s take the rule of *reductio ad absurdum* in the following form: if we hypothetically accept **A** and, within the scope of the hypothesis, we conclude not-**A**, then we may accept not-**A** outside the scope of the hypothesis. I don’t think that the rule of *reductio* possesses more obviousness than (T). Say we are facing a theoretical difficulty and, in order to overcome it, we need either to reject that rule or to reject (T). It is not clearly preferable to

⁴ Traditional emotivists and prescriptivists were happy to say things like ‘Murder is evil’ but deny that such sentences are true. Were they incoherent? If their semantic views on ethical terms were correct, there was no incoherence. For if a sentence such as ‘Murder is evil’ is a command or manifests an emotion, like an interjection, they did not perform an assertion when uttering it. It is not possible to assert a command or interjection, although it is of course possible to say it in earnest. If, on the other hand, their semantic views were incorrect, and their ethical utterances were assertions, then their position may be considered incoherent.

reject (T). After all, we can replace the rule of *reductio* with the following weaker one: if we hypothetically accept **A** and, within the scope of the hypothesis, we conclude not-**A**, then we should not accept **A** outside the scope of the hypothesis.

The extent to which a principle is obvious is a matter of judgement. So an argument against classical logic, however intuitive it may be, will not be able to convince everyone. Still, it can convince those who find that accepting its assumptions is more compelling than accepting classical logic in its entirety. Also, when we judge classical logic, there are important considerations other than obviousness, such as its simplicity and familiarity. So intuitive arguments against classical logic are not the end of the matter. Still, they may detract from the power of those considerations.

Abstracting away from many details, the basic idea of the arguments that follow is that, relying on classical logic, we can derive contradictions not only from extensional principles, such as the schema ‘**S** is true iff p ’, but also from intensional semantic claims, such as the statement ‘The word “true” expresses the concept of being true’ and the schema ‘**S** means that p ’. Those claims are harder to reject than their extensional counterparts, so deviating from classical logic becomes correspondingly more appealing.

1. The First Argument against Classical Logic

Suppose there is such a property as truth, the property of being true. Then, if we reason in accordance with classical logic, we can infer a contradiction from some premisses that are quite obvious, including two semantic principles. The reasoning concerns the sentence (L). I shall also name it ‘(α)’ and compare it with (β), which is the sentence

(L) is true.

The first semantic principle is that the truth of a negation amounts to the untruth of what is negated:

(N) For any sentences **S** and **S'** such that **S** is a negation of **S'**, **S** is true iff **S'** is not true.

I say that **S** is a *negation of S'* iff **S'** means just that p and **S** means just that not- p . For if **S'** and **S** mean that p and that not- p respectively, but

either of them has additional meanings, there arise complications that need not detain us here. As a consequence of (N), (α) is true iff (β) is not true.

The second semantic principle will say in effect that a predication $\lceil \mathbf{t}$ is $\mathbf{F} \rceil$ is true iff the referent of the term \mathbf{t} has the property expressed by the predicate \mathbf{F} . For example, the sentence ‘Socrates is wise’ is true just in case the referent of ‘Socrates’ has the property expressed by ‘wise’. Some care, however, is required here if we are to sidestep problems that result from empty singular terms and from predicates that express no property. Unless there are such things as mythological characters, the term ‘Zeus’ has no referent. Yet, intuitively, the sentence ‘Zeus is an ancient god’ is true. Again, a version of Russell’s paradox arguably shows that there is no such property as non-self-exemplification. In other words, there is no such thing as the property of being a property that does not exemplify itself. If so, the predicate ‘property that does not exemplify itself’ expresses no property. Yet the sentence ‘Wisdom is a property that does not exemplify itself’ is true.

Thus the second semantic principle becomes:

- (P) For any nonempty singular term \mathbf{t} and any predicate \mathbf{F} that expresses a property, the sentence $\lceil \mathbf{t}$ is $\mathbf{F} \rceil$ is true iff the referent of \mathbf{t} has the property expressed by \mathbf{F} .

I think that, in this form, the principle is acceptable. Names \mathbf{t} that have more than one bearer (e.g. ‘London’) and predicates \mathbf{F} that, because of ambiguity, express more than one property (e.g. ‘horse race’) are not counterexamples. In their case, the right-hand side of the biconditional fails, since there is no such thing as the referent of \mathbf{t} or there is no such thing as the property expressed by \mathbf{F} . But the left-hand side also fails. $\lceil \mathbf{t}$ is $\mathbf{F} \rceil$ is not true *simpliciter*, although it may be true on this or that one of its readings.⁵

The predicate ‘true’ expresses the property of being true. The only way in which that might not be the case is for there to be no such property as the property of being true. If there is, as we have here supposed, ‘true’ certainly expresses it. That is so just as, if there is such a property as wisdom, ‘wise’ expresses it, and if there is such a property as self-exemplification, ‘exemplifies itself’ expresses it. Hence, by (P),

⁵ For the same reason, indexical predicates \mathbf{F} , like ‘friend of mine’, are not counterexamples.

the sentence ‘(L) is true’ is true iff the referent of ‘(L)’ has the property expressed by ‘true’. But ‘true’ expresses no property other than truth. So the property of being true is *the* property expressed by ‘true’. And the referent of ‘(L)’ is sentence (α) . Therefore, sentence (β) is true iff (α) has the property of being true.

The following schema about exemplification of properties is very intuitive:

- (E) If there is such a property as the property of being F , then for everything $x(x$ is F iff x has the property of being F).

For instance, if there is such a property as being red, then anything is red just in case it has the property of being red. So if there is such a property as being true, then (α) is true iff it has the property of being true. Thus, since we have assumed the existence of truth, (α) is true iff it has the property of being true.

It is now a matter of transitivity and symmetry for biconditionals to infer that (β) is true iff it is not true. (And it is a matter of transitivity, symmetry and contraposition for biconditionals to infer that (α) is true iff it is not true.) Hence, by classical logic, (β) is both true and not true — contradiction.

This argument, in which we did not use schema (T), can be read as a *reductio* of the assumption that there is such a property as the property of being true. That is not an advisable response, though. For the argument can be rephrased as one about the concept of being true. Instead of talking about expressing a property, we shall talk about expressing a concept, and instead of saying that something has a property, we shall say that something comes under a concept. If the argument is rephrased, one cannot hold to the same response. Denying that there is such a concept as the concept of being true is even more difficult than denying that there is such a property as the property of being true. One can use the paradoxes to argue that the notion of truth is in some sense defective, but one can hardly argue that it does not exist.⁶

Of course, we sometimes talk about concepts of truth, in the plural, or say that the notion so-and-so is *a* concept of truth. According to

⁶ For the distinction between properties and concepts see, e.g., H. Putnam’s [15, pp. 429–431]. Properties and concepts are individuated differently. Since water is the substance H_2O , it is natural to consider that the property of containing water is the same as the property of containing H_2O . Still, the concept of containing water is other than the concept of containing H_2O .

M. Dummett [3, pp. 165–166], a notion is a concept of truth provided one can say that an assertion is correct (by one’s preferred standard) just in case the asserted statement falls under that notion. Then, the concept of admitting in principle of a proof and even the concept of having been proved count as notions of truth. Such concepts, however, are simply variants of the notion that the word ‘true’ expresses in ordinary language when we talk about the truth of statements, beliefs, speakers’ words, etc. That notion can be described as *the* concept of truth, despite the presence of its variants.

So, setting aside the idea that there is no such property as truth, the contradiction shows a conflict between some semantic or property-theoretic claims and classical logic. The claims include (N), (P) and (E), and staunch supporters of classical logic will probably target the first two. Things are not so simple, though. We know from the liar and kindred paradoxes that certain semantic principles lead to contradictions when combined with standard logic. The recalcitrant principles are disquotational. One of them is (T). Another one is the schema

$$\text{For everything } x, x \text{ satisfies } \mathbf{G} \text{ iff } x \text{ is } F, \quad (\text{S})$$

where the letter ‘ F ’ is to be replaced by a monadic predicate while ‘ \mathbf{G} ’ is to be replaced by a name of that predicate. By putting the predicate *predicate that does not satisfy itself* as F and its quotation as \mathbf{G} and by also letting the predicate instantiate the universal quantification, we infer that ‘predicate that does not satisfy itself’ satisfies ‘predicate that does not satisfy itself’ iff it is a predicate that does not satisfy itself. In other words, the predicate satisfies itself iff it does not. That is of course a version of Grelling’s paradox. On the other hand, principles (N), (P) and (E) are not disquotational. So they are not the kind of claim that we know from experience may be involved in getting us into trouble.⁷

It may be replied that (N) and (P) are not innocent, since they are similar to principles that allow us to derive, in a standard Tarskian way, biconditionals stating the truth-conditions of negations of atomic

⁷ That may also help with the question whether the analogue of (E) for concepts should be accepted for defective concepts too. If a notion is defective, that may be a reason for not applying it to some things. But I think that unless it is a reason for not using the notion at all, it can hardly be a reason for excluding it from a principle, such as (E) or its analogue for concepts, which is not disquotational. In any case, I do not think that the notion of truth is defective.

sentences, including the problematic biconditional ‘(L) is true iff (L) is not true’. Indeed, one variant of (P) is

- (P′) For any nonempty singular term \mathbf{t} and any monadic predicate \mathbf{F} , the sentence $\lceil \mathbf{t}$ is $\mathbf{F} \rceil$ is true iff the referent of \mathbf{t} satisfies \mathbf{F} .

With the help of (N) and (P′), we can get up to the following biconditional: (L) is true iff it is not the case that the referent of ‘(L)’ satisfies the predicate ‘true’. In order to conclude that (L) is true iff it is not and thus reach contradiction, we must also use some disquotational premisses. In particular, we need a premiss to the effect that, for everything x , x satisfies ‘true’ iff x is true or, at least, one to the effect that (L) satisfies ‘true’ iff it is true. If one is unwilling to accept that the fault lies with classical logic, it is reasonable to consider a premiss of either kind to be what causes the problem. For it instantiates the schema (S), which we know to be involved in paradox. Indeed, the claim ‘For everything x , x satisfies “true” iff x is true’ is similar to ‘For everything x , x satisfies “predicate that does not satisfy itself” iff x is a predicate that does not satisfy itself’ in that both are instances of (S) in which the letter ‘ F ’ has been replaced with a semantic predicate. So, whether or not the fault lies with classical logic, (N) and (P′) are not to blame for the contradiction. This indicates that (N) and (P) are not to blame for the contradiction in the present argument.

Of course, we did invoke a disquotational premiss in deriving this contradiction. We accepted that if there is such a property as being true, then ‘true’ expresses the property of being true. That premiss can hardly be considered suspect. The claim that ‘true’ expresses the property (or the concept) of being true is more secure than the claim that everything satisfies ‘true’ iff it is true. If need be, one could stick to the former claim but deny the latter. (But then, presumably one should also reject the principle ‘Everything satisfies “true” iff it has the property of being true’. For, given the existence of truth, that principle, together with (E), implies that everything satisfies ‘true’ iff it is true.) The friends of classical logic might perhaps suggest that our reasoning should be seen as a *reductio* of the premiss in question: there is such a property as being true, but the predicate ‘true’ does not express it. I think that a response of this kind is no less heroic than abandoning classical logic.⁸ Indeed, in my judgement, rather than go so far in making

⁸ That does not mean that no such response has existed. According to N. Smith’s

bold claims outside the realm of logic, it is preferable to deviate from classical logic.⁹

2. The Second Argument against Classical Logic

The next argument to the effect that (T) is not to blame for the contradiction is that in some variants of the liar we end up in contradiction without having invoked (T) or any other schema, rule or statement about truth. Instead of (T), we invoke some even more obvious schema about meaning or saying. This possibility is not frequently touched upon in the literature.¹⁰ Admittedly, the relevant variants of the liar usually concern sentences in a symbolic or half-symbolic idiom that poses problems independent of the paradox and goes beyond the linguistic apparatus of both natural language and the symbolic language of first-order logic.

See the sentence

$$\forall p[\text{if } (L') \text{ means that } p, \text{ then } \neg p]. \quad (L')$$

(L') involves quantification into sentence position: both occurrences of the letter 'p' within the brackets are in syntactic positions in which a sentence can be written and not in positions in which a name can be written. It is not clear what is the sense of quantifying into sentence position.

[18, p. 182], when we utter (L), either we do not refer to the sentence we utter or we do not say of what we refer to that it is not true.

⁹ In a variant of the argument, we replace (N) and (P) with

(N') For any sentences **S** and **S'** such that **S** is a negation of **S'**, **S** is true iff **S'** is false, and

(P'') For any nonempty singular term **t** and any predicate **F** that expresses a property, the sentence $\ulcorner \mathbf{t} \text{ is } \mathbf{F} \urcorner$ is false iff the referent of **t** lacks the property expressed by **F**

and also rely on the schema

(E') If there is such a property as the property of being *F*, then for everything *x* $[(x \text{ is not } F) \text{ iff } x \text{ lacks the property of being } F]$.

So we infer that (α) is true iff it is not true. Those who believe that no statement is neither true nor false are likely to accept (N), whereas those who believe that some statements are neither true nor false are likely to endorse (N') instead. The conditional is material throughout.

¹⁰ But see A. Prior's [14, pp. 140–141], S. Haack's [8, p. 150] and P. Pagin's [12]. Also see S. Read's [16, pp. 13–15], I. Rumfitt's [17, pp. 41–45] and J. Stern's [21, pp. 56–62, in particular the references on p. 56].

It may not even be clear whether it has a sense. Such quantification is not present in first-order logic, and at least *prima facie* it does not seem to be present in natural language either. One might argue that e.g. the sentence ‘ $\forall p[p \text{ or not-}p]$ ’ is meaningless and we do not understand it and, if some are under the illusion that they do, this is happening because it is similar to other sentences that are meaningful but do not involve quantification into sentence position (such as the sentence ‘For every proposition x , x is true or x is not true’). Moreover, (L) contains the intensional ‘means that’. The intensionality is shown by the fact that arguments of the form ‘The sentence so-and-so means that q ; q iff r ; hence, the sentence so-and-so means that r ’ are generally invalid. For example, it is invalid to infer from ‘The sentence “Gras ist grün” means that grass is green’ and ‘Grass is green iff it possesses chlorophyll’ to ‘The sentence “Gras ist grün” means that grass possesses chlorophyll’. As is well-known, intensional contexts present various semantic and logical problems.

I take it that quantification into sentence position has a sense. However, it should not be understood as quantification over some objects. There is no kind of entities such that sentences of the form ‘ $\forall p \dots$ ’ say how all entities of that kind are, and sentences of the form ‘ $\exists p \dots$ ’ say how at least some entities of that kind are. For if we take the sentences ‘ $\forall p \dots$ ’ and give them paraphrases of the form ‘It is the case for every entity x of the kind so-and-so that $\dots x \dots$ ’, or if we take the sentences ‘ $\exists p \dots$ ’ and give them paraphrases of the form ‘There is an entity x of the kind so-and-so such that $\dots x \dots$ ’, either the paraphrases will be meaningless or we shall introduce concepts that are not expressed in the original sentences. The paraphrases will lack meaning if we write the variable ‘ x ’ in sentence position without combining it with any predicate. That is our predicament when, for example, we paraphrase the sentence ‘ $\forall p[p \text{ or not-}p]$ ’ as ‘For every proposition x , either x or it is not the case that x ’. What is meaningless does not have the same meaning as the original sentence. Alternatively, new concepts will appear in the predicates that combine in sentence position with the variable ‘ x ’. When, for instance, we paraphrase ‘ $\forall p[p \text{ or not-}p]$ ’ as ‘For every proposition x , x is true or x is not true’, we introduce the concept of truth, which does not seem to be involved in the original sentence. But if a paraphrase does not express the concepts expressed in the original sentence and only those, then it does not have the same linguistic meaning as that sentence.¹¹

¹¹ D. Grover, J. Camp and N. Belnap would consider that the concept of truth

Also, quantification into sentence position may but need not be understood substitutionally. A substitutional reading selects a set of sentences as the substitution class for the letters ‘ p ’, ‘ q ’, etc. Then, ‘ $\lceil \forall p \mathbf{A} \rceil$ ’ is true iff, for every member of the substitution class, the sentence that results from \mathbf{A} by replacing ‘ p ’ with that member is true, whereas ‘ $\lceil \exists p \mathbf{A} \rceil$ ’ is true iff the same condition is met by at least one member of the substitution class. So the generality of the quantification is restricted to the substitution class. If, for example, the class contains no sentence meaning that the Earth rotates, then whether ‘ $\forall p[p \text{ or not-}p]$ ’ is true does not depend on whether it is the case that the Earth rotates or the Earth does not rotate.

It seems, however, that we can understand ‘ $\forall p[p \text{ or not-}p]$ ’ in such a way that, although it is not a quantification over propositions, its truth requires every disjunctive proposition that is a case of excluded middle to be true. That reading of ‘ $\forall p[p \text{ or not-}p]$ ’ is not substitutional. For if it were, there would be a substitution class, selected implicitly, which ensured that level of generality. In fact, any set of sentences that we might be implicitly selecting as the substitution class is inappropriate. No set of sentences from the natural or symbolic languages we use allows such generality, since presumably there are propositions that do not find expression in any sentence of those languages. If a language possesses no word or phrase expressing a concept, no sentence in the language expresses any proposition involving that concept. There must be concepts, that is, ways of thinking about things, which we do not grasp and which are not expressed in our languages. Indeed, it would be boastful to think that we shall ever enrich our idioms to such an extent that they express every concept and the set of their sentences allows the required level of generality. So we can understand quantification into sentence position in a way that is not substitutional.

We can interpret ‘ $\exists p$ ’ as meaning ‘ $\neg \forall p \neg$ ’. And we can interpret the sentence ‘ $\forall p[p \text{ or not-}p]$ ’ by considering that it says generally and universally just what the sentence ‘The Earth rotates or the Earth does

is involved in the sentence ‘ $\forall p[p \text{ or not-}p]$ ’, for according to the theory they present in [7] the phrase ‘it is true’ can function just like the two occurrences of ‘ p ’ within the brackets. They also consider that ‘ $\forall p[p \text{ or not-}p]$ ’ admits of an English paraphrase of the form ‘For every proposition, ...’. Nevertheless, they would agree with my conclusion that quantification into sentence position should not be understood as quantification over propositions or any other objects. Indeed, they even extend that point to the English paraphrases ‘For every proposition, ...’; see especially p. 114.

not rotate' says about whether the Earth rotates (or what 'There are aquatic mammals or there are no aquatic mammals' says about whether there are aquatic mammals) and by also considering that it is entirely general, so that its generality is not restricted to what can be said in any particular languages. The former point implies that there is no conceptual difference between 'The Earth rotates or the Earth does not rotate' and ' $\forall p[p \text{ or not-}p]$ ' except for the replacement of the particular topic about the Earth with the universal quantification. Although the issue may be controversial, I for one accept that (L'), and more generally quantification into sentence position, has a sense along the lines just suggested. That sense is not objectual, since it introduces no extraneous concepts (such as the notion of truth) when we move from the particular to the universal, but it is not substitutional either.¹²

In the case of (L'), the reasoning that leads to contradiction presupposes not (T) but the following schema:

$$\mathbf{S} \text{ means that } p, \tag{M}$$

where the letter ' p ' is to be replaced with a declarative sentence while ' S ' is to be replaced with a name of that sentence. So one instance of (M) is this:

$$(\mathbf{L}') \text{ means that } \forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p].$$

Indeed, sentence (L') is not ambiguous. It means only that $\forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$.

Suppose that $\forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$. By universal instantiation we conclude that if (L') means that $\forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$, then $\neg \forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$. Hence, by *modus ponens*, $\neg \forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$.

Consequently, by *reductio*, $\neg \forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$. Hence, $\exists p[(\mathbf{L}') \text{ means that } p, \text{ and } p]$. But the sentence (L') means *only* that $\forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$. Therefore, $\forall p[\text{if } (\mathbf{L}') \text{ means that } p, \text{ then } \neg p]$ — contradiction. (The last step in the reasoning is analogous to the step from the premisses ' $\exists x[\text{George loves } x, \text{ and } x \text{ is a cat}]$ ' and 'George loves only Felix' to the conclusion 'Felix is a cat'.)¹³

¹² See T. Williamson's [26, pp. 257–264] on quantification into sentence position that is neither objectual nor substitutional.

¹³ The reasoning bears some similarity to the paradox involving quotation and presented by Tarski in [22, pp. 161–162] but avoids the special difficulties of quantify-

It is true that the schema (M) should be accepted subject to some restrictions. First, we should not apply it to meaningless sentences. If we apply it to the sentence ‘Letrins burb’, we shall espouse something meaningless, namely,

‘Letrins burb’ means that letrins burb.

This is meaningless because its last two words lack sense. Second, it is wrong to apply (M) to sentences that contain indexicals. Indexicals introduce special problems. If I apply (M) to the sentence ‘I am a philosopher’, I shall say,

‘I am a philosopher’ means that I am a philosopher.

Yet the sentence does not mean that I am a philosopher. Nor is there any other person about whom it means that that person is a philosopher. The sentence is a type used by various people. My utterances of it may mean that I am a philosopher, and George’s utterances of it may mean

ing into quotation. If our quantifiers have the sense I suggested, it is not possible to quantify into quotation. To see that, let ‘quote’ be an operator that takes a sentence and yields a singular term; when the sentence is closed, the term refers to it, but otherwise the term lacks reference. So ‘quote(the Earth is round)’ refers to the sentence ‘The Earth is round’, whereas ‘quote(p)’ does not refer to the letter ‘ p ’. Quantifying into quotation, if at all possible, should be easier with ‘quote’ than it is with ordinary quotation marks. Let c be the sentence ‘Aeneas betrayed Dido’. When, for example, we write ‘ $\forall p$ quote(p) = c ’ and the quantifier is the kind I suggested, do we quantify into the context ‘quote(...)?’ If we do, then ‘ $\forall p$ quote(p) = c ’ says universally just what ‘quote(the Earth is round) = c ’ says about the roundness of the Earth, what

(1) quote(Aeneas betrayed Dido) = c

says about Aeneas’ betrayal of Dido, what

(2) quote(Dido was betrayed by Aeneas) = c

says about Dido’s betrayal by Aeneas, etc. This presupposes that what (1) says about Aeneas’ betrayal of Dido is what (2) says about Dido’s betrayal by Aeneas. But then, (1) and (2) say the same thing; there is no difference in content between them. For the issue of Aeneas’ betrayal of Dido is the same as the issue of Dido’s betrayal by Aeneas. In fact, (1) and (2) cannot have the same content, as they are not even materially equivalent. (1) is true, while (2) is false. Trying to quantify into quotation in the sense I suggested amounts to trying to say generally just what various sentences say about a number of respective topics, when in fact those sentences do not say the same thing about those topics. So it is an attempt at doing the impossible.

that he is a philosopher, but the linguistic meaning of the sentence-type is not connected with any particular person.¹⁴

There is no problem with applying (M) to ambiguous sentences. For example, we may say,

‘Horse races differ widely’ means that horse races differ widely.

This instance of the schema is ambiguous because the clause introduced by ‘that’ has two meanings. The subject of ‘means’ is not ambiguous: it is a name of a certain string of words. The instance of (M) is true on both of its readings. For ‘Horse races differ widely’ means both that races between horses differ and that races of horses differ. So the ambiguity is harmless. On the other hand, we should not say,

‘Horse races differ widely’ means only that horse races differ widely.

For this is false on both of its readings.

The restrictions on (M) do not concern sentence (L’). I accepted it as meaningful, and at least if we fix the language by interpreting ‘means’ as ‘means in ordinary or technical English’, the sentence does not appear to involve any indexicals. Apart from those restrictions, (M) seems to be beyond dispute. Can we doubt that ‘Snow is black’ means that snow is black, or that ‘Whales are mammals’ means that whales are mammals? There are certainly semantic and logical differences between, on the one hand, the statements

‘Snow is black’ means that snow is black

and

$$\begin{array}{l} \forall p[\text{if } (L') \text{ means that } p, \text{ then } \neg p] \text{ means that} \\ \forall p[\text{if } (L') \text{ means that } p, \text{ then } \neg p] \end{array} ,$$

and, on the other, the statements

‘Snow is black’ means ‘Snow is black’

and

$$\begin{array}{l} \forall p[\text{if } (L') \text{ means that } p, \text{ then } \neg p] \text{ means} \\ \forall p[\text{if } (L') \text{ means that } p, \text{ then } \neg p]. \end{array}$$

It is not here necessary to dwell on the differences. For all the differences, I think that the former statements are as indisputable as the latter.

¹⁴ The term ‘indexicals’ here extends to all expressions some of whose semantic features vary with the context of utterance while their meaning remains fixed.

If we endorse (T) without restricting it to nonparadoxical sentences, we can accept that each instance of that schema is entailed by an instance of (M).

S means that snow is black

entails

S is true iff snow is black.

For no sentence could have meant that snow is black without being true iff snow is black. If, on the other hand, one is unwilling to abandon classical logic, one is tempted to deny that (M) implies (T) in the case of paradoxical sentences. The idea is to accept, for example, that

(L) means that (L) is not true

but deny that

(L) is true iff (L) is not true.

That option is tempting because (M) is more difficult to doubt than (T). However, the case of (L') shows that such a move does not suffice to uphold (M).¹⁵

We occasionally justify a homophonic instance of (T) (that is, an instance in which the name that replaces 'S' is a quotation of the sentence used on the right-hand side) by invoking the homophonic instance of (M) for the same sentence. We say things like 'The condition that is necessary and sufficient for the sentence "Snow is black" to be true is that snow is black, since that is what the sentence means'. So Williamson [24, p. 149] writes, 'given that "TW is thin" means that TW is thin, what more could it take for "TW is thin" to be true than for TW to be thin?' This occurs in the discussion of an objection to biconditionals such as '“TW is thin” is true iff TW is thin'. The objection is that if 'TW is thin' is neither true nor false, then calling it 'true' is false, so according to standard three-valued logics the biconditional is neither true nor false. As far as I know, we never attempt to justify a homophonic instance of (M). This indicates that (M) is indeed more difficult to doubt than (T).

¹⁵ If one assents to '(L) means that (L) is not true', or rather to the stronger '(L) means only that (L) is not true', but dissents from '(L) is true iff (L) is not true', one should reject the definition of '... is true' as ' $\exists p[\dots \text{ means that } p, \text{ and } p]$ '. Since '(L) means only that (L) is not true' entails ' $\exists p[(L) \text{ means that } p, \text{ and } p] \text{ iff } (L) \text{ is not true}$ ', the definition in question validates the inference from the former sentence to '(L) is true iff (L) is not true'. For definitions of sentential truth along those lines, see W. Kühne's [11, pp. 346–349].

Of course, in order to end up in contradiction, we did not rely merely on the relevant instance of schema (M), that is, on the claim that the sentence (L') means that $\forall p$ [if (L') means that p , then $\neg p$]. We also presupposed that (L') means *only* that. We presupposed that (L') does not have two or more meanings. But this assumption, too, seems to me to be innocent. Apart from the machinery of quantification, the words occurring in (L') are unambiguous. Also, the sentence does not admit of more than one syntactic analysis. As for the quantification into sentence position, it can be understood substitutionally and so differently from the way I understand it. But this ambiguity can be ignored, since it can easily be removed. I could introduce a new symbol instead of '∀', and this would certainly have no more than one meaning.¹⁶

Also, in order to end up in contradiction, we made some inferences. Some of them, like *modus ponens*, are validated by classical logic. Others involve quantification into sentence position, but are analogous to

¹⁶ Williamson, in [25], postulates that semantic terms, such as 'true', 'false' and 'say', shift in meaning over time, as we use them in reasoning. His approach, however, does not seem to be an adequate response to the paradoxes. Such a shift can explain away the contradiction in a biconditional like '(L) is true iff (L) is not true' in two ways: either the meaning that the relevant sentence has on the right-hand side is not the meaning under which it is evaluated on the left-hand side (so there is no reason for asserting the biconditional) or a semantic term, e.g. 'true', has a different meaning on the two sides (so there is no problem with logic). Now, whether or not a declarative sentence x has more than one reading (meaning) and whether or not it contains indexicals, we can normally assert,

S is true on its present reading in the current context iff p ,

where ' p ' has been replaced with that sentence and 'S' with a name of that sentence. We are entitled to assert that biconditional if our utterance of it will take place in a single context (as it normally will) and involve a single reading of x . For the meaning and context under which we evaluate x when we utter the left-hand side will be the same as the meaning it has, and the context in which it is used, when we utter the right-hand side. But then, take the sentence

(L) is not true on its present reading in the current context. (L)

(L) is true on its present reading in the current context iff (L) is not true on its present reading in the current context. When I uttered that biconditional, the meaning under which (L) was evaluated on the left-hand side was what the 'its present reading' on that side referred to. And this can hardly be other than the meaning that the sentence had on the single temporally close occasion where it was used and not mentioned, that is, when I uttered the right-hand side. Again, even if 'true' and 'reading' change meaning over time, it is hard to believe that they changed meaning in the course of uttering the biconditional. So postulating an ambiguity in those words or just in (L) does not help to rid us of the contradiction.

inferences that are validated by classical logic and involve quantification into name position instead. For example, classical logic allows us to make inferences of the form ‘ $\neg\forall x[\text{if } A \text{ then } \neg B]$; hence $\exists x[A \text{ and } B]$ ’. Similarly, we here made an inference of the form ‘ $\neg\forall p[\text{if } A \text{ then } \neg B]$; hence $\exists p[A \text{ and } B]$ ’. Finally, we reasoned from ‘ $\exists p[a \text{ means that } p, \text{ and } \dots p \dots]$ ’ and ‘ $a \text{ means only that } X$ ’ to ‘ $\dots X \dots$ ’. Reasoning of this form is analogous to inferences (such as the one about Felix) that involve the word ‘only’ and are standardly transcribed into the language of first-order logic in such a way that classical logic validates the transcriptions.

The first inference on the way to contradiction consisted in instantiating the letter ‘ p ’ with (L'). That inference might be problematic if the quantification were substitutional. For it would presuppose that (L') belonged to the substitution class, and, on a reading of substitutional quantification into sentence position, it is illegitimate for the substitution class to contain the quantified sentences; if it does, the sentences lack truth-conditions. When substitutional quantification is read metalinguistically, there is no lack of truth-conditions. We stipulate that e.g. ‘ $\forall p[p \text{ or not-}p]$ ’ is true iff, for every member of the substitution class C , the sentence that results from ‘ $p \text{ or not-}p$ ’ by replacing ‘ p ’ with that member is true. By means of that stipulation, we specify a metalinguistic condition that is necessary and sufficient for the truth of ‘ $\forall p[p \text{ or not-}p]$ ’. The condition (that all the instances of ‘ $\forall p[p \text{ or not-}p]$ ’ should be true) is assigned to the quantified sentence irrespective of whether ‘ $\forall p[p \text{ or not-}p]$ ’ belongs to C and whether its instances have already been assigned truth-conditions. Any instance that still has no truth-condition is not true. When, on the other hand, substitutional quantification is not read metalinguistically, then, although the stipulation about ‘ $\forall p[p \text{ or not-}p]$ ’ is the same, what we do by means of it is different. We presuppose that the instances of ‘ $\forall p[p \text{ or not-}p]$ ’ already have truth-conditions and specify that what is necessary and sufficient for the truth of the quantified sentence is the joint satisfaction of those conditions. Thus, ‘ $\forall p[p \text{ or not-}p]$ ’ is assigned a truth-condition only if its instances have already been assigned truth-conditions. If the quantified sentence belongs to C , then one of its instances is ‘ $\forall p[p \text{ or not-}p]$ or not- $\forall p[p \text{ or not-}p]$ ’, which does not acquire a truth-condition before ‘ $\forall p[p \text{ or not-}p]$ ’. So if the quantified sentence belongs to C , it is assigned no truth-condition.¹⁷

¹⁷ Substitutional quantification is read metalinguistically in Haack’s [8, pp. 42 and 49–50] and P. Horwich’s [9, pp. 4, 25 and 32–33]. The metalinguistic reading is

At any rate, instantiating ‘ p ’ with a quantified sentence is not problematic if quantification into sentence position is understood in the way adumbrated above. The sense of ‘ $\forall p[p \text{ or not-}p]$ ’, for example, was specified without presupposing that all its instances had already been assigned truth-conditions. So the specification of its sense does not fail because ‘ $\forall p[p \text{ or not-}p]$ or not- $\forall p[p \text{ or not-}p]$ ’ is accepted as an instance. In this respect, ‘ $\forall p[p \text{ or not-}p]$ ’ is like ‘Every disjunctive proposition that is a case of excluded middle is true’. We can explain the latter sentence without assuming that all its instances, that is, all sentences of the form ‘The proposition that \dots or not- \dots is true’, have antecedently acquired truth-conditions. So the explanation does not fail if the result of putting ‘every disjunctive proposition that is a case of excluded middle is true’ in place of the dots had no truth-condition before the explanation. There is no vicious circularity threatening the legitimacy of the move from ‘ $\forall p[\dots p \text{ ---}]$ ’ to ‘ $\dots \forall p[\dots p \text{ ---}]$ ’. On the contrary, the generality of the quantification in the former sentence justifies the move. So I submit that whoever supports classical logic and accepts that quantification into sentence position has a sense along the lines I suggested earlier cannot but approve the inferences that were made here.

To sum up, if quantifying into sentence position has a sense along the lines adumbrated, the schema (M) is beyond dispute (apart from the restrictions I mentioned), and there is no room for supporting the classical logic of quantification but not accepting the analogous patterns of reasoning that we followed in handling quantification into sentence position, then the liar and its variants should be tackled by deviating from classical logic. Indeed, the crux of the matter, which may be obscured by issues about quantification that are not clear at first sight, is that there is a conflict between (M) and classical logic. And even if it is acceptable to abandon (T) in order to keep classical logic in every subject matter, abandoning (M) is too high a price.

explicitly rejected in M. Dunn and N. Belnap’s [4, pp. 184–185] and S. Soames’s [19, pp. 91–92]. As Soames points out (p. 91), the two readings are not strictly equivalent. R. Barcan Marcus, in [1, pp. 82–85], and S. Kripke, in [10, pp. 329–332 and 367–368], argue that it is illegitimate for the substitution class to contain the quantified sentences.

Acknowledgements. This paper is a revised version of my talk at a conference in honour of Mark Sainsbury which took place in Barcelona in October 2012. Mark had been my PhD supervisor, and later on a colleague, at King’s College London. I learned a very great deal from him both about philosophy and about teaching. For all of us who were his students, his work was the starting-point for our own investigations. A Greek version was read at the seminar of the research project Thales – UOA – APrePoSMa in my department (December 2013). The second argument in the paper had also been presented at the staff seminar in my department, at the seminar of the graduate programme “Logic and the Theory of Algorithms and Computation” in the University of Athens and at a conference on paradox and logical revision in the Munich Center for Mathematical Philosophy. On all those occasions I benefited from the remarks that were made. I also benefited from some referee reports, particularly one from *LLP*.

References

- [1] Barcan, M. R., “Quantification and ontology”, pages 75–87 in her *Modalities: Philosophical Essays*, New York: Oxford University Press, 1993 (reprinted with changes from *Noûs* 6, 3 (1972): 240–250).
- [2] Barwise, J., and J. Etchemendy, *The Liar: An Essay on Truth and Circularity*, New York: Oxford University Press, 1987.
- [3] Dummett, M., *The Logical Basis of Metaphysics*, London: Duckworth, 1991.
- [4] Dunn, J., and N. Belnap, “The substitution interpretation of the quantifiers”, *Noûs* 2, 2 (1968): 177–185. DOI: [10.2307/2214704](https://doi.org/10.2307/2214704)
- [5] Field, H., *Saving Truth from Paradox*, Oxford: Oxford University Press, 2008. DOI: [10.1093/acprof:oso/9780199230747.001.0001](https://doi.org/10.1093/acprof:oso/9780199230747.001.0001)
- [6] Gaifman, H., “Pointers to truth”, *The Journal of Philosophy* 89, 5 (1992): 223–261. DOI: [10.2307/2027167](https://doi.org/10.2307/2027167)
- [7] Grover, D., J. Camp and N. Belnap, “A prosentential theory of truth”, *Philosophical Studies* 27, 2 (1975): 73–125. DOI: [10.1007/BF01209340](https://doi.org/10.1007/BF01209340)
- [8] Haack, S., *Philosophy of Logics*, Cambridge: Cambridge University Press, 1978. DOI: [10.1017/CBO9780511812866](https://doi.org/10.1017/CBO9780511812866)
- [9] Horwich, P., *Truth*, 2nd edition, Oxford: Clarendon Press, 1998. DOI: [10.1093/0198752237.001.0001](https://doi.org/10.1093/0198752237.001.0001)

- [10] Kripke, S., “Is there a problem about substitutional quantification?”, pages 325–419 in G. Evans and J. McDowell (eds.), *Truth and Meaning: Essays in Semantics*, Oxford: Oxford University Press, 1976.
- [11] Künne, W., *Conceptions of Truth*, Oxford: Clarendon Press, 2003. DOI: [10.1093/0199241317.001.0001](https://doi.org/10.1093/0199241317.001.0001)
- [12] Pagin, P., “‘Means that’ and the liar”, pages 203–230 in H. Malmgren (ed.), *Kvantifikator för en Dag: Essays Dedicated to Dag Westerståhl on His Sixtieth Birthday*, Philosophical Communications, Web Series, no. 35, Sweden: Göteborg University, Department of Philosophy, 2006.
- [13] Priest, G., *In Contradiction: A Study of the Transconsistent*, 2nd edition, New York: Oxford University Press, 2006. DOI: [10.1093/acprof:oso/9780199263301.001.0001](https://doi.org/10.1093/acprof:oso/9780199263301.001.0001)
- [14] Prior, A., “Some problems of self-reference in John Buridan”, pages 130–146 in his *Papers in Logic and Ethics*, London: Duckworth, 1976.
- [15] Putnam, H., “The nature of mental states”, pages 429–440 in his *Mind, Language and Reality: Philosophical Papers, vol. 2*, Cambridge: Cambridge University Press, 1975 (first published as “Psychological predicates”, pages 37–48 in W. Capitan and D. Merrill (eds.), *Art, Mind, and Religion*, Pittsburgh: University of Pittsburgh Press, 1967). DOI: [10.1017/CB09780511625251.023](https://doi.org/10.1017/CB09780511625251.023)
- [16] Read, S., “The truth-schema and the liar”, pages 3–17 in S. Rahman, T. Tulenheimo and E. Genot (eds.), *Unity, Truth and the Liar: The Modern Relevance of Medieval Solutions to the Liar Paradox*, Logic, Epistemology, and the Unity of Science, vol. 8, Springer, 2008. DOI: [10.1007/978-1-4020-8468-3_1](https://doi.org/10.1007/978-1-4020-8468-3_1)
- [17] Rumfitt, I., “Truth and meaning”, *Aristotelian Society Supplementary Volume* 88 (2014): 21–55.
- [18] Smith, N., “Semantic regularity and the liar paradox”, *The Monist* 89, 1 (2006): 178–202. DOI: [10.5840/monist200689139](https://doi.org/10.5840/monist200689139)
- [19] Soames, S., *Understanding Truth*, New York: Oxford University Press, 1999. DOI: [10.1093/0195123352.001.0001](https://doi.org/10.1093/0195123352.001.0001)
- [20] Stephanou, Y., “A propositional theory of truth”, *Notre Dame Journal of Formal Logic* 59, 4 (2018): 503–545. DOI: [10.1215/00294527-2018-0013](https://doi.org/10.1215/00294527-2018-0013)
- [21] Stern, J., *Toward Predicate Approaches to Modality*, Cham, Switzerland: Springer, 2016. DOI: [10.1007/978-3-319-22557-9](https://doi.org/10.1007/978-3-319-22557-9)
- [22] Tarski, A., “The concept of truth in formalized languages”, pages 152–278 in his *Logic, Semantics, Metamathematics*, translated by J. H. Woodger and edited by J. Corcoran, 2nd edition, Indianapolis: Hackett, 1983.

- [23] Van Fraassen, B., “Presupposition, implication, and self-reference”, *The Journal of Philosophy* 65, 5 (1968): 136–152. DOI: [10.2307/2024557](https://doi.org/10.2307/2024557)
- [24] Williamson, T., “Vagueness and ignorance”, *Aristotelian Society Supplementary Volume* 66 (1992): 145–162. DOI: [10.1093/aristoteliansupp/66.1.145](https://doi.org/10.1093/aristoteliansupp/66.1.145)
- [25] Williamson, T., “Indefinite extensibility”, *Grazer Philosophische Studien* 55 (1998): 1–24. DOI: [10.5840/gps19985512](https://doi.org/10.5840/gps19985512)
- [26] Williamson, T., “Truthmakers and the converse Barcan formula”, *Dialectica* 53, 3–4 (1999): 253–270. DOI: [10.1111/j.1746-8361.1999.tb00186.x](https://doi.org/10.1111/j.1746-8361.1999.tb00186.x)
- [27] Yablo, S., “New grounds for naive truth theory”, pages 312–330 in J. C. Beall (ed.), *Liars and Heaps: New Essays on Paradox*, Oxford: Clarendon Press, 2003.
- [28] Zardini, E., “Truth without contra(di)ction”, *The Review of Symbolic Logic* 4, 4 (2011): 498–535. DOI: [10.1017/S1755020311000177](https://doi.org/10.1017/S1755020311000177)

YANNIS STEPHANOU

Department of History and Philosophy of Science
National and Kapodistrian University of Athens
Ilissia Campus, 157 71 Athens, Greece
ystephan@phs.uoa.gr