PARADOXES OF TRUTH AND MODALITY*

Volker Halbach

1st April 2012

1 TRUTH IN PHILOSOPHY

Modern Western philosophy is highly compartmentalised: Philosophers spezialise in ethics, metaphysics, epistemology, philosophy of mind or logic. Some philosophers working on problems in logic have specialised in the theory of semantic paradoxes; the subject has the reputation of being inaccessible, very formal and mind-boggling. Hence most philosophers working in areas outside logic – for instance, in epistemology or ethics – are only too happy to leave the paradoxes to these specialists. Division of labour seems to relieve most philosophers of the nasty task to bother about the liar and related paradoxes.

Unfortunately division of labour and compartmentalisation doesn't work always. In particular, also philosophers outside logic should be worried about the problems arising from the paradoxes. Logic and the theory of paradoxes are so deeply interwoven with our philosophical theories that they cannot easily put aside and left to the logicians.

In this talk I will introduce the semantic paradoxes and related paradoxes concerning modal notions; and I will explain why these paradoxes may affect wide parts of philosophy and thus why the average philosophers should bother about the paradoxes.

2 THE LIAR PARADOX

From all the paradoxes the liar paradox is the most famous. It is not clear who stated the paradox first. Eubulides of Miletus, who lived in the 4th century BC, is often cited as the first author to mention the paradox. Based on a passage in the bible by Paul, some people claim that Epimenides invented the liar paradox. However, there seem to be references to the liar paradox in much earlier work in the 5th century BC in the Indian literature, and perhaps also in the Mohist literature.

To state the paradox, I introduce some terminology. To talk about certain sentences I will label them. For instance, the symbol ($\diamond$) can be used to label a sentence:

($\diamond$) Shanghai is in China.

*I thank Shao Qiangjin for his encouragement and help that made my trip to China and this lecture possible.

Then obviously the sentence labelled ⋄ is the sentence "Shanghai is in China". It is also obvious that the sentence labelled ⋄ is true.

Now we consider another sentence with (†) as its label:

(†) The sentence labelled † is not true.

This sentence is self-referential in the sense that it says something about the sentence labelled with the symbol †, and that sentence is not another sentence but rather the sentence itself: the cross symbol † in the sentence is used to refer to the sentence that has † as a label.

Sentences such as † are known as *liar sentences*.

To see why this sentence is paradoxical, Assume the following:

(1) The sentence labelled † is not true.

Now look at the sentence with the label †; it's the same sentence as 1. So if the assumption 1 is true, then the sentence † is true. But this contradicts our assumption 1, which says that † is *not* true. Therefore our assumption 1 is false and we must reject it and conclude the opposite:

(2) The sentence labelled † is true.

Now let's look at the sentence that is claimed to be true. What does the sentence labelled † say? It says that the sentence labelled † is *not* true thereby contradicting our conclusion 2. So it cannot be true – contrary to what we said before in 2. This is a contradiction: We have shown that the sentence labelled † cannot be not true. But we have also shown that it cannot be true. This is the liar paradox.

You might wonder and ask: What is so bad about contradictions? To demonstrate how disastrous contradictions are, I will prove that Shanghai is in Brazil. This is false of course, but using a contradiction one can prove *anything*, including the claim that Shanghai is in Brazil.

To see how I arrive at this conclusion, consider again the liar sentence † from above:

(†) The sentence labelled † is not true.

Let's not assume anything at the outset. We know that

> Either † is true or † is not true.

But we don't know whether it is true or not, but only that one of the alternatives must hold. So also one of the following *three* alternatives must hold:

> Either † is true or † is not true or Shanghai is in Brazil.

Now consider the first alternative: Assume that † is true. That means that the sentence

> The sentence labelled † is not true

is true, because this is the sentence †. So the first alternative cannot obtain.

May be the second alternative obtains? So assume the second alternative:

> † is not true.

If sentence labelled † – that is, the sentence "The sentence labelled † is not true" – is not true, then the sentence † must be true. So against our assumption † *is* true and the second alternative fails as well.

Now we have eliminated the first and the second alternative from

> Either † is true or † is not true or Shanghai is in Brazil.

and only the last alternative is left: Shanghai is in Brazil. We have derived a fairly surprising geographic claim by purely truth-theoretic reasoning. By the same pattern of reasoning it could also be argued that Shanghai is not in Brazil. Then contradiction would have spread to geography. It should be clear that there is nothing particular to the claim the Shanghai is in Brazil. As I said, using a contradiction one can argue for any claim. Contradiction is contagious. It spreads like a disease.

Now somebody doing geography could say: "I don't accept your conclusion. This is just a cheap trick. Truth is a philosophical notion and I just don't use it." I think that would be a reasonable response, even though even geographers occasionally use the word "true". But presumably geographers could live without the word "true" and similar notions; it is not essential to doing geography.

But the philosopher cannot get away so easily. Philosophers need to use the word "true". This is not only because truth is a topic in itself in philosophy but also because truth is needed in many areas of philosophy.

Consider epistemology for instance. The epistemologist cannot just shrug his shoulders and ignore truth. He cannot easily stop using the truth predicate and avoid the derivation of unwanted consequences by no longer using the word "true". For the epistemologist needs the notion of truth.

Here is an example of how it is needed. According to an old story, going back to Plato at least, a person *knows* something if and only if the person believes it with good reasons and it is true. There is an extensive literature about whether this account of knowledge as true justified belief is tenable – in particular, the so-called Gettier problem seems to pose a problem – but I don't intend to discuss whether this story correct or not. I would just like to emphasise that without truth we cannot even *state* the thesis. Truth is needed to make a claim that is about knowledge and many philosophers have felt that truth serves here only the purpose of allowing us to make a generalised claim. I agree, but if truth is threatened by paradox, then the account of knowledge as true justified belief is also at risk.

Metaphysics is another area where the word "true" is used essentially. Realism is another philosophical doctrine that is often stated with the use of the truth predicate. Realists claim against their opponents that there is an independent reality. It is not easy to spell out precisely what that means, but one way to describe the realist position is the following: According to the realist, there are *truths* that can never be verified, that is, there are truths no one might ever have good reasons to believe. The existence of this possibly unverifiable truths has replaced the somewhat unclear characterisation using an "independent reality". So we have a thesis from metaphysics that cannot even be stated without the concept of truth (or some related notion). It would be odd if the viability of realism depended on the liar paradox, but it seems many realists need the truth predicate and therefore they are dependent on a solution of the paradoxes.

The discussion about realism reappears in many different forms. In moral philosophy, for instance, some authors claim that moral claims like "Stealing is wrong" are neither true nor false. I don't intend to take sides in this discussion either; I only would like to emphasise that without

the word "true" the entire discussion does not get off the ground. We cannot even start to discuss moral anti-realism if we don't have a truth predicate available.

I hope these examples convince you that truth is needed in philosophy. It permeates many philosophical discussion not necessarily as a notion that is at the centre of the discussion, but as a notion that is needed – perhaps sometimes in a way similar to logical expressions such as "and", "all" and similar expression; they may not be in themselves at the front stage of a discussion, but without them we wouldn't be able to discuss many philosophical issues at all.

3 SELF-REFERENCE

So far it might not have become clear why it is *truth* that is causing the problem. It could be conjectured that something else, something that is dispensable and less important than truth is causing the trouble. In particular, one might suspect that in the liar sentence

(†) The sentence labelled † is not true.

it is not the word "true" that spells trouble but rather the odd way in which the sentence makes a claim about itself.

Labelling a sentence in a certain way and then using that very label – in our case the symbol † – in the labelled sentence looks suspicious. If we don't allow this, one might hope that we can avoid the paradox. In particular, we might hope that the real source of the paradox is not the notion of truth but rather a method of attaching labels to sentences in a funny way.

But the paradox cannot be so easily resolved. The paradox can be regained in a way where the label is not used in the sentence it labels. Liar sentences can be formulated without using a label such as † that is then used in the labelled sentence. For instance, one can also obtain a paradox in the following way:

> *The sentence in italics on this page is not true.*

That will still look suspicious. That the phrase "the sentence in italics" refers back to itself seems as weird as the method involving the label †. We may want to say that this is not legitimate. We don't seem to use such ways of talking about sentences in many cases. Generally one might suspect that the paradoxes can be easily solved by disallowing sentences that refer back to themselves as in the examples above.

If such sentence are to be avoided, one can obtain a paradox in a slightly less direct way. Imagine a postcard that has written on one side

> The sentence on the other side is not true.

When you turn the postcard around and look at the other page you read the following sentence:

> The sentence on the other side is true.

Note that on the second side the "not" is missing. It is not hard to see that what is written on the postcard yields a paradox similar to the liar paradox. This time no labels are needed. An each sentence on either side of the postcard doesn't make a claim about itself; it merely says something about the sentence on the other side of the postcard.

It is not easy to get rid of the paradoxes. One would not only have to disallow the use of labels such as † and the use of descriptions such as "the sentence in italics", our basic methods for

talking about expressions would have to be given up. When doing linguistics, logic, computer science and other subjects, we talk about expressions or, in other words, we reason syntactically. This involves talking about replacing certain expressions with other expressions and the like. But I won't go into details here, because the details are somewhat complicated.

One doesn't even need the theory of syntax and talk about expressions. Gödel (1931) showed that one can assign numbers to expressions, so that each expression has a unique number. He also showed that one can obtain effects similar to the one at work in liar sentences simply by doing basic arithmetic. Again I can't into details here. But all these observations have convinced philosophers that one cannot avoid paradoxes by disallowing the use of labels like † in a sentence that contains this label.

To see that it is really truth rather than the use of strange labelling and the like, one can try to replace truth with other notions in liar sentence. For instance, if in † the word "true" is replaced with some other notion not prone to paradox, then the problem doesn't arise. For instance, the sentence

(†) The sentence labelled † is not an English sentence.

is just false and by no means paradoxical. Or you can replace "is not true" with "is not a sentence in Wu Chinese" and you'll get an unproblematic true sentence.

This shows that truth is to blame for the paradox not the way that sentences make claims about themselves.

4 PROPOSITIONS

Some philosophers would say that my reasoning so far is based on a mistake. They would say that what is true are not sentences but rather what is expressed by sentences. Most philosophers would call the things that are expressed by sentences *propositions*. So these philosophers would say that the proposition expresses by the sentence "Shanghai is in China" is true but that the sentence itself is not true.

If that view is adopted, the paradox seems to disappear. What is expressed by the sentence † is then true, because it says that the sentence † is not true, and that is a correct claim because no sentence is true; only *propositions* are true.

However, the paradox returns. Consider the following sentence

(P) The sentence labelled P does not express a truth proposition.

It is not hard to see that a contradiction can be obtained even if one believes that only propositions can be true.

It also doesn't help to say that P doesn't express a proposition at all, because if the sentence P doesn't express a true proposition at all, then it doesn't express a true proposition; and that is what is claimed in P.

I conclude that is doesn't help to use propositions as objects that can be true. Propositions may be useful for many purposes; but they won't help us to solve the paradox.

5 MEDDLING WITH LOGIC

I have argued that it is not a sensible strategy to give up truth; we need it in philosophy. We also cannot block the paradox by putting a ban on sentences that make a claim about themselves.

If we blocked the paradox in this way, we would have to give up many other unproblematic things as well. So truth is to be blamed for the contradiction and its consequences. I think most philosophers in the field would follow me up to this point. But now we have reached the point where disagreement begins.

Some philosophers would say that the best way to deal with the paradoxes is to change our logic. The basic ways of reasoning that allow us to derive absurd consequences by using paradoxes have to be curtailed. Some philosophers would go even further an claim that in my reasoning leading to the absurd conclusion that Shanghai is in Brazil I made a mistake. Some of the steps in my derivation are incorrect and I used the wrong logic, they would protest. What is called classical logic has never been the correct logic.

There are many ways to mutilate classical classical.

Some influential philosophers propose to replace classical logic with what is called *paraconsistent* logic.[1] They propose to change the logic in such a way that we can have contradiction without being able to derive absurd consequences such as the claim that Shanghai is in Brazil. Hence they can say that the liar sentence † is true and that † is not true, but that still Shanghai is not in Brazil.

Other philosophers, among them Horsten (2011), reject contradictions and block the derivation of absurd consequences very early on. They would reject claims like:

> Either † is true or † is not true.

Usually these philosophers have to reject even sentences like

> If † is true, then † is true.

This has lead them to come up with new baroque logics that avoids the derivation of absurd consequences such as the claim that Shanghai is in Brazil. In particular, many authors now work on new ways of understanding "if" that avoids paradox but doesn't force them to reject sentences like the one above. The most prominent author who has many proposals in this direction is Field (2008). But as these proposals become very technically complicated, I won't explain them here.

Many advocates of solutions of the paradoxes that involve nonclassical logic, play down the costs of the change from classical to nonclassical logic. But the notion of truth is so deeply entrenched in philosophy that I don't think that one will get away with making only small negligible changes to our philosophical theories. If we really give up classical logic, I expect that we have to use it not only in connection with truth but also in connection with other notions that are closely related. For instance if truth is defined in terms of true justified belief, then presumably the nonclassical behaviour of the truth predicate is transmitted to knowledge as well. And from there it will spread to other areas of philosophy. If we opt for a different logic, we will have to revise large areas of philosophy and subject it to the new logic.

As I prefer to avoid the rewriting of so many philosophy books, I will look at other ways to avoid the bad consequences of the liar paradox.

[1] When Priest (1987) began to advocate his view, it was considered esoteric and bordering on the absurd. But over the years his account of the paradoxes called *dialetheism* according to which a sentence can be true together with its negation has become very popular especially in Australia from where it has spread.

6 AXIOMATIC APPROACHES TO TRUTH

In the derivation of the liar paradox and the claim that Shanghai is in Brazil I made use of certain assumptions on the word "true". For instance, I invoked an assumption that allowed me to go from

> If the sentence "The sentence labelled † is not true" is true

to the sentence

> The sentence labelled † is not true.

This step is justified if we have generally the schema

> If "*A*" is true, then *A*

where *A* is a declarative sentence (that is a sentence that is used to make an assertion). I also need the reverse direction

> If *A*, then "*A*" is true.

Plausible as these claims may look, I think we should reject them. In fact Tarski (1935) argued in the most influential essay on truth, that we have to give up the two schemata above. This again is not without a price. We will have to see how useful a truth predicate with these weaker properties is. But we will be able to retain classical logic, which is in my view a huge advantage.

Various philosophers have looked at the question of which assumption we can consistently make about truth and then what we can do with a truth predicate that satisfies these assumptions. The discussion becomes very technical and here I cannot go any further into it. I recommend the book by Horsten (2011) as a starting point at least for those who have mastered a good logic course.

7 MODAL PARADOX

So far I have concentrated on the liar paradox. As I said, it's the most famous paradox. But there are many more paradoxes. Some of them involve truth again, but others apply to other notions. There are many varieties, but I'll give only a paradox about necessity that is very similar to the liar paradox and that is known as *Montague's paradox*, because the basic structure of the argument can be found in Montague (1963).

We begin with a sentence that is very similar to a liar sentence but with "necessary" in the place of "true".

(M) M isn't necessary.

This line is short for the "The sentence labelled M is not necessary" or – if you prefer to ascribe necessity to propositions – "The proposition expressed by M is not necessary."

We can now reason as follows:

> If M is necessary, then M is necessary.

This is just a triviality. But because M is the sentence "M isn't necessary" we also have the following:

If M is necessary, then "M isn't necessary" is necessary.

Now the general principle "If "*A*" is necessary, then *A*" is applied to obtain the following line:

If M is necessary, then M isn't necessary.

Hence we have established M:

M isn't necessary.

It cannot be otherwise. So we conclude that "M isn't necessary" is necessary and therefore that

M is necessary.

This contradicts the previous line.

Therefore the notion of necessity is threatened by inconsistency as well. The notions of aprioricity, analyticity and knowability fall prey to analogous paradoxes.

8 PERNICIOUS INTERACTIONS

Many philosophers who don't want to give up classical logic say that, in order to be safe from the paradoxes, they don't make any assumption on how the truth predicate behaves when it's applied to sentences containing the truth predicate. This is very restrictive of course. We don't have

If "*A*" is true, then *A*

for all sentence *A* anymore, but only for those sentences *A* that do not contain the word "true". Although this strategy limits what we can do with truth, it seems to be a safe strategy. Here I will not evaluate this strategy, but instead I'll argue that despite the first initial impression of safety, disaster is lurking behind the corner.

The strategy is general. If the strategy is adopted for truth, it seems reasonable to apply this strategy also to other notions such as necessity.

So the proposal is that we don't make any assumptions on whether sentences containing "true" are true or not; and we don't make any assumptions on whether sentences containing "necessary" are necessary or not. As I said, this may be very restrictive, but it seems a safe way to block the paradoxes. In fact in this way we can block the liar paradox and Montague's paradox.

Although now our ways of reasoning about truth and necessity are very limited, the paradox still strike. Our assumptions about necessity and truth lead to a contradiction even if we don't apply truth to sentences containing "true" and necessity not to sentences containing "necessary".[2]

To facilitate the derivation of the paradox, I use the name N for the sentence

(N) ♡ is necessary.

The sentence ♡ is defined in the following way:

(♡) N is not true.

To derive a contradiction, we begin with a triviality:

[2] I have presented the following paradox in (2006).

1. If N is true, then N is true.

2. If ♡ is necessary, then N is true. (*This follows from 1 and the definition of N and the assumption "If "A" is true, then A" for sentences A without "true"*.)

3. If " N is not true" is necessary, N is not true. (*This is an instance of the principle "If "A" is necessary, then A" applied to a sentence A not containing "necessary"*.)

4. If ♡ is necessary, N is not true. (*This is obtained from 3, because " N is not true" has been defined as* ♡.)

5. ♡ is not necessary. (*This follows from 2 and 4.*)

6. N is not true. (*This is because N says that* ♡ *is necessary.*)

7. ♡ is necessary. (*In the previous line we have established that N cannot be true. So* ♡ *necessary.*)

Lines 5 and 7 contradict each other. Thus we have derived a paradox. In no line have we applied "is necessary" to a sentence that contains the word "necessary"; and we haven't applied "is true" to any sentence containing "true".

It is possible to block the liar paradox by making no assumptions about the truth of sentences containing "true"; it's also possible to block Montague's paradox about necessity by making no assumptions about the necessity of sentences containing "necessary". But as soon as we try to combine both solutions, a new paradox emerges.

This paradox provides another illustration of my remark I made at the beginning: compartmentalisation is a risky strategy: even if solutions work for truth and necessity separately, there is no guarantee that the solutions can be consistently combined. Compartmentalisation is dangerous.

9 CONCLUSION

I could go on to present further paradoxes. For instance, Horsten and Leitgeb (2001) showed that combining assumptions about future and past leads to the paradoxical result that there is no future. There are many paradoxes that arise from the interaction of predicates such as "is true", "is necessary", "is analytic", "is known", "is a priori" and further predicates.

But being able to treat these notions in one common framework is important for philosophy. Philosophers are discussing all the time how all these notions are related. For instance, Kant and many later philosophers famously claimed that there are a priori truths that are not analytic. I don't intend to take sides in this discussion about Kant's doctrine, but if we don't have a consistent logical framework in which we can express such claims without running into paradoxes, then also philosophers who are not directly interested in the paradoxes should be worried. They cannot simply leave the solution of the paradoxes to the logicians.

As we have seen, logicians have proposed drastic measures; some of them have proposed to give classical logic. Even if that solves the problem with paradox, other philosophers – metaphysicians, epistemologists and so on – should think twice before they adopt this strategy, because they will have to reason in the new nonclassical logic also about their own topics.

Metaphysicians and epistemologists can also not just hope that they can use separate solutions of the liar paradox and of the other paradoxes that are proposed by logicians. What metaphysicians and epistemologists need are solutions that still work when one is using truth, necessity and so on together in one's reasoning. What we need are solutions that will allow us to reason using all these notions together.

As a logician I admit that we are still not close to having such an integrated theory. At best we have separate solutions for each notion, that is, truth, necessity and so on. But there are many paradoxes and riddles out there that arise from the interaction of modal notions and truth that have not even been discovered, let alone be solved.

References

Hartry Field. *Saving Truth From Paradox*. Oxford University Press, Oxford, 2008.

Kurt Gödel. über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte Fur Mathematik*, 38-38:173–198, 1931.

Volker Halbach. How not to state the T-sentences. *Analysis*, 66:276–280, 2006. Correction of printing error in vol. 67, 268.

Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, 2011.

Leon Horsten. *The Tarskian Turn: Deflationism and Axiomatic Truth*. MIT Press, Cambridge, Mass., 2011.

Leon Horsten and Hannes Leitgeb. No future. *Journal of Philosophical Logic*, 30:259–265, 2001.

Richard Montague. Syntactical treatments of modality, with corollaries on reflexion principles and finite axiomatizability. *Acta Philosophica Fennica*, 16:153–67, 1963. Reprinted in (Montague, 1974, 286–302).

Richard Montague. *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, New Haven and London, 1974. Edited and with an introduction by Richmond H. Thomason.

Graham Priest. *In Contradiction*. Martinus Nijhoff, Dordrecht, 1987. second expanded edition Oxford University Press 2006.

Alfred Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica Commentarii Societatis Philosophicae Polonorum*, 1:261–405, 1935. reprinted as 'The Concept of Truth in Formalized Languages' in (Tarski, 1956, 152–278); page references are given for the translation.

Alfred Tarski. *Logic, Semantics, Metamathematics*. Clarendon Press, Oxford, 1956.

Volker Halbach
New College, Oxford University
OX1 3BN Oxford
England