

PHILOSOPHY OF LOGIC AND LANGUAGE

WEEK 3: THE LIAR PARADOX

JONNY MCINTOSH

OVERVIEW

Last week:

- What does Tarski achieve with his truth definitions?
- Is his approach to the Liar Paradox too restrictive?
- Kripke's alternative approach to the Liar Paradox

This week:

- The Simple Liar
- Self and circular reference
- Denying bivalence
- Allowing true contradictions
- Contextualist approaches

THE SIMPLE LIAR

First, assume the following identity:

1. $\lambda = \text{'}\lambda \text{ is false'}$

We then have, as an instance of the T-schema:

2. λ is true IFF λ is false

Also, by the principle of **BIVALENCE**:

3. Either λ is true or λ is false

We can then reason as follows:

4. Assume λ is true
5. From 2., it then follows that λ is also false
6. Assume instead that λ is false
7. From 2., it then follows that λ is also true

Since either way we have it that λ is both true and false, it follows from 3. that:

8. λ is true and λ is false

Now 8. is not yet an explicit contradiction — i.e. a sentence of the form, P and not P. But given a plausible assumption about falsity and negation, one quickly follows.

The plausible assumption is that:
9. A sentence is false IFF its negation is true

I'll leave the rest of the proof as homework!

DENYING 1.

We might try to deny 1., the assumption that λ is the sentence ' λ is false'.

We can then hang on to the conception of truth, embodied by the T-schema, that delivers 2.

That is, we can accept each instance obtained from the following T-schema by replacing 'X' with the name of a sentence and 'P' by that sentence:

'X' is true IFF P

We can also hang on to the principles of classical logic, including **BIVALENCE**, which deliver 3., and validate the subsequent reasoning.

But what exactly is wrong with assuming that λ is the sentence ' λ is false'?

Tarski thinks that the problem is that it ignores the distinction between different levels in a hierarchy of object- and meta-languages, or of truth predicates.

But:

- Tarski's approach is very restrictive.
- It is unclear that it has any independent motivation.

Are there any other grounds for denying 1.?

One thought: λ refers to itself.
Perhaps the paradox is a result of **SELF** reference?

But we can derive the paradox without self reference:

- $\lambda_A = \text{'}\lambda_B \text{ is true'}$
- $\lambda_B = \text{'}\lambda_A \text{ is false'}$

Another thought: λ_A and λ_B refer to each other.
Perhaps the problem is **CIRCULAR** reference?

But arguably the paradox can be derived without even circular reference:

- $\lambda_0 = \text{'Each } \lambda_n \text{ is false, where } n > 0 \text{'}$
- $\lambda_1 = \text{'Each } \lambda_n \text{ is false, where } n > 1 \text{'}$
- $\lambda_i = \text{'Each } \lambda_n \text{ is false, where } n > i \text{'}$

If λ_0 is true, then all "later" λ_i are false.
So λ_1 is false. In which case some "later" λ_j is true.
So λ_0 must be false, and some "later" λ_j true.

We have shown that λ_0 must be false. But the reasoning can be repeated to show that each "later" λ_i is false as well.

If λ_1 is true, then all "later" λ_i are false.
So λ_2 is false. In which case some "later" λ_i is true.
So λ_1 must be false, and some "later" λ_i true.

Stephen Yablo argues that this shows that λ_0 is also true. In which case we have a contradiction that doesn't even involve circular-reference.

DENYING BIVALENCE

If we accept the assumption that $\lambda = \text{'}\lambda \text{ is false'}$, what else can we do?

One option is to give up **BIVALENCE** — roughly, that every sentence is either true or false — and so give up 3.

This is a common thought. We saw a precise working out of it last week: Kripke's theory.

But does this really get to the heart of the problem?

First, assume the following identity:

1. $\lambda = \text{'}\lambda \text{ is not true'}$

Next, as an instance of the T-schema:

2. λ is true IFF λ is false

Also, by the **LAW OF EXCLUDED MIDDLE**:

3'. Either λ is true or λ is not true

We can then reason as follows:

4. Assume λ is true
5. From 2., it then follows that λ is also not true.
6. Assume instead that λ is not true
7. From 2., it then follows that λ is also true.

Since either way we have it that λ is both true and not true, it follows from 3' that:

8. λ is both true and not true.

This is the problem of **REVENGE**: given a solution to one version of the Liar Paradox, it seems possible to construct a new, **STRENGTHENED LIAR**.

In fact, it is not so clear that this is a problem for Kripke. But there is another problem in the vicinity, namely that his solution is **INEFFABLE**.

Here's why the **STRENGTHENED LIAR** is not a problem for Kripke:

- Neither ' λ is true' nor ' λ is not true' come out as true on Kripke's theory.
- So the relevant instance of the Law of Excluded Middle, 3', cannot be asserted.

Here's why there is nevertheless a problem of **INEFFABILITY**:

- Kripke wants to say that λ is not true.
- But the sentence he uses to say this is ' λ is not true'!
- So how can he truly say what he wants to say?

The problem: Kripke's solution to the Liar Paradox cannot be stated within the language for which his truth predicate is defined — on pain of paradox!

Kripke was aware of the problem:

"If we think of the minimal fixed point, say under the Kleene valuation, as giving a model of natural language, then the sense in which we can say, in natural language, that a Liar sentence is not true must be thought of as associated with some later stage in the development of natural language, one in which speakers reflect on the generation process leading to the minimal fixed point. It is not itself a part of that process. The necessity to ascend to a metalanguage may be one of the weaknesses of the present theory. The ghost of Tarski's hierarchy is still with us."

Does the problem matter? Not obviously. Perhaps it is enough that we can avoid the paradox; we don't also need to be able to say *how* it is avoided.

As we will see, **CONTEXTUALISTS** take the problem very seriously. And offer a solution.

DENYING EXPLOSION

Kripke takes a **PARACOMPLETE** approach to the Liar Paradox, allowing that not all sentences of the form, $P \vee \neg P$, are true.

Others take a **PARACONSISTENT** approach to the Liar Paradox, allowing that some sentences of the form, $P \wedge \neg P$, are true.

This approach is also known as **DIALETHEISM**, and true sentences of the form, $P \wedge \neg P$, as **DIALETHEIAS**.

On this approach, the reasoning that takes us to a contradiction is not just valid, but **SOUND**. All it shows is that certain contradictions are *true*.

Like Tarski and Kripke, proponents of this approach can hang on to the conception of truth embodied by the T-schema.

Unlike Kripke, they can also hang on to the classical **LAW OF EXCLUDED MIDDLE**. And no other step in the reasoning that leads to contradiction need be jettisoned.

So what's the problem? What, if anything, is wrong with allowing for true contradictions?

PRINCIPLE OF EXPLOSION: from a contradiction, one may deduce anything at all.

This is also part of classical logic. Unless it is rejected, paraconsistency will allow us to infer anything we like from the conclusion of the Liar Paradox — e.g. that $1=0$.

To deal with this, paraconsistent theorists, like Graeme Priest, develop logics in which **EXPLOSION** is abandoned.

The key is the idea that, where Kripke allows for truth value **GAPS**, sentences that are *neither* true *nor* false, paraconsistent logics allow for truth value **GLUTS**, sentences that are *both* true *and* false.

An advantage over paracomplete approaches:

- It is possible to define truth in such a way that the status of λ as both true and false can be stated within the object language.

An open question:

- Is it possible to define truth in such a way that the status of non-defective sentences as, e.g., *merely* true can be stated in the object language?

Another issue with :

- How can the paraconsistency theorist understand **DISAGREEMENT**?

If one person says 'A' and another '¬A', the paraconsistency theorist allows that both may be right. Similarly if one says 'A is true' and the other says 'A is false' or 'A is not true'.

So how are we supposed to capture the fact that they may be disagreeing?

One option: distinguish the mental state of belief that ¬A from the mental state of **REFUSAL** to believe that A, and characterise disagreement in terms of the latter.

(For students studying Ethics: there are interesting parallels here with the moves made by expressivists in metaethics in response to the Frege-Geach problem.)

CONTEXTUALISM

Suppose that I am in Oxford and talking on Skype to Mike, who is in California. I say 'It is 6pm here'. Mike says 'It is not 6pm here'. We both speak truly.

Now suppose I offer the following argument:

- It is 6pm here
- It is not 6pm here
- So, it is 6pm here and it is not 6pm here

Obviously what has gone wrong is that I have overlooked the **CONTEXT-SENSITIVITY** of the sentences I use — the fact that they express different propositions, or otherwise have different semantic statuses, in different contexts of utterance.

Similarly, according to **CONTEXTUALIST** approaches to the Liar Paradox, λ expresses different propositions, or otherwise has different semantic statuses, in different contexts of utterance.

Suppose that λ is the sentence 'The sentence written on the board in room 5 does not express a true proposition', and is the only sentence written on the board in room 5. And suppose I say:

The sentence written on the board in room 5 does not express a true proposition.

My statement — made here, in room 6 — seems perfectly coherent, and in the envisaged circumstances, true.

The contextualist thus promises not just a way of blocking the paradox, but also of solving the **REVENGE** problems that confront other solutions. The problem had two steps:

- First, we try to resist the paradox by saying that λ is in some way defective.

- Second, we conclude from this first step that λ must after all be true.

Contextualists promise a way of giving both of these thoughts their due — without leading to paradox.

- λ is in some way defective, and in initial contexts doesn't express a truth.
 - This blocks the reasoning that leads to paradox.
- But when we use λ to explain how, we shift to a new context in which it *does* express a truth.

Different contextualists spell out the idea in different ways.

Tyler Burge employs the notion of a Tarskian hierarchy of truth predicates, each with a different subscript, which is silent, invisible, and supplied by context.

- In initial contexts, the silent subscript is, say, i . In this context, λ is defective.
- When we use λ to say that it is defective, however, we move to a context in which the silent subscript is rather some $k > i$.

Charles Parsons employs the idea of a **QUANTIFIER DOMAIN RESTRICTION**: a contextually supplied restriction on the domain of things we are talking about.

To see the idea, suppose I threw a party for people studying or teaching philosophy in Oxford, and tell you:

- Every student came to the party

You might reply by pointing out that no engineering students came to the party, and say:

- Not every student came to the party

What's happened? Roughly: in the context in which I spoke, the domain of things being talked about contained only philosophy students, but in the context in which you spoke, the domain had been expanded to include all Oxford students.

Parsons' thought is that a sentence S is true, as uttered in a context C , IFF there is a proposition P that S expresses in C , and P is true.

The idea is then that in initial contexts, the domain of quantification does not contain any proposition expressed by λ .

But in contexts in which we explain what has gone wrong, the domain *does* include a proposition expressed by λ , and this proposition is true.

Questions for Burge and Parsons:

- Why think that truth predicates are context-sensitive in the ways they suggest?
- What exactly is the mechanism by which context sets the subscript or domain restriction?

It is also unclear that the **REVENGE** problem is adequately addressed. Consider:

- 'This sentence is not true at any level'
- 'This sentence is not true in any context'

Proponents of contextualism can (and do) try to argue that quantification over all levels or contexts is impossible, e.g. on the grounds that there are no absolutely unrestricted quantifiers.

SUMMARY

We've looked at:

- The Simple Liar
- The role of self- and circular-reference
- Attempts to avoid the problem by denying **BIVALENCE**

And also:

- Problems of **REVENGE** and **INEFFABILITY**
- Attempts to avoid the problem by allowing true contradictions
- Attempts to avoid the problem by appeal to context-sensitivity

Next week:

- Model-theoretic conceptions of logical consequence