PHILOSOPHY OF LOGIC AND LANGUAGE

WEEK 6: THE LIAR PARADOX

JONNY MCINTOSH

INTRODUCTION

Last week, we examined the question of what, exactly, Tarski achieved by means of his famous definitions of truth.

They don't seem to provide the basis of an **DEFINITION** of the concept of truth, but perhaps they provide the basis of an **EXPLICATION** of it.

And in any case, they certainly provide a way of talking about the true sentences of certain languages without running risk of the Liar Paradox.

Yet as we'll see, it's a very restrictive approach. We'll look at some alternatives. But first, a reminder of how the Liar goes.

THE SIMPLE LIAR

First, assume the following identity:

1. λ = 'λ is false'

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

2. 'λ is false' is true IFF λ is false

We can then reason as follows:

3. Assume λ is true
4. From 1., it follows that 'λ is false' is true
5. From 2., it then follows that λ is false

6. And if we assume instead that λ is false
7. From 2., it follows that 'λ is false' is true
8. From 1., it then follows that λ is true

We can therefore conclude:

9. λ is true IFF λ is false

But now, by the principle of **BIVALENCE**:

10. Either λ is true or λ is false

It therefore follows that:

11. λ is both true and false

Other than classical logic, the assumptions are:

- Premise 1: λ = 'λ is false'
- The **T-SCHEMA**: 'λ is false' is true IFF λ is false
- **BIVALENCE**: either λ is true or λ is false

DENYING 1

Tarski's solution is to insist on the distinction between object- and meta-language, and deny 1. But his approach is very restrictive.

In place of Tarski's hierarchy of *languages*, we might try to introduce a hierarchy of restricted **PREDICATES**:

- 'true$_0$' (roughly: 'is a true sentence not containing any truth predicate')
- 'true$_1$' ('is a true sentence whose only truth predicates are "true$_0$"')
- 'true$_2$' ('is a true sentence whose only truth predicates are "true$_0$" or "true$_1$"')
- and so on...

This way, we can distinguish a hierarchy of syntactically individuated levels *within* a single language.

But even this is very restrictive, as is nicely brought out by Kripke's Watergate examples.

PROBLEM 1: How do we determine the appropriate subscript for truth predicates?

John Dean: 'Nothing Nixon said about Watergate up to the time of his resignation was true'

To assign a subscript, we would need to know the highest level to which sentences uttered by Nixon belonged.

PROBLEM 2: In some cases, it is not even possible in principle to assign a subscript.

John Dean: 'Most of Nixon's Watergate-related statements are not true'

Nixon: 'Most of John Dean's Watergate-related statements are true'

The subscript on Dean's 'true' will have to be higher than any subscript on any truth predicate uttered by Nixon.

But so too must the subscript on Nixon's 'true' be higher than any on any truth predicate uttered by John Dean.

The subscripting approach treats this pair as paradoxical. But it's perfectly conceivable that both sentences are true!

Suppose neither uttered any other sentence containing the word 'true'. And that 90% of the sentences Nixon uttered were false, while 90% of those Dean uttered were true.

There is a **RISK** of paradox, however.

Suppose that, apart from these, Nixon and Dean both uttered an even number of sentences, exactly half of which are true and half false.

The lesson: strategies that employ **SYNTACTIC** criteria to screen off paradoxical sentences will rule out sentences for which there is only a **RISK** of paradox.

DENYING BIVALENCE

If we accept the assumption that λ = 'λ 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. Kripke's own theory is an attempt to work it out precisely. But does it really get to the heart of the problem?

First, assume the following identity:

1. λ = 'λ is not true'

Next, as an instance of the T-schema:

2. 'λ is not true' is true IFF λ is not true

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 1 and 2., it then follows that λ is also not true.
6. Assume instead that λ is not true
7. From 1 and 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. For in fact, neither 'λ is true' nor 'λ is not true' are true on his theory.

So the relevant instance of the Law of Excluded Middle, premise 3 in the foregoing argument, cannot be asserted.

But there is another problem in the vicinity for Kripke, namely that his solution is **INEFFABLE**.

- Kripke wants to say that λ is not true.
- But the sentence he uses to say this is λ!
- 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." (Kripke 1975, p. 714)

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. First, a different approach...

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 in the Liar Paradox 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.

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:

- 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 contextualists:

- 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 limitations of Tarski's approach
- 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