

# PHILOSOPHY OF LOGIC AND LANGUAGE

## WEEK 2: TARSKI AND KRIPKE

JONNY MCINTOSH

## OVERVIEW

Last week:

- The Liar Paradox as a motivation for Tarski's work on truth
- Formal and material constraints on acceptable definitions
- Tarski's techniques: recursion and variable assignments

This 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

## TARSKI'S ACHIEVEMENT?

## DEFINING TRUTH

Did Tarski show how to define the concept of truth, i.e. define predicates that have the same meaning as the English predicate 'true'?

One problem: 'true' seems to apply to **PROPOSITIONS** (what we use sentences to say) where Tarski's predicates apply to **SENTENCES**.

The question here concerns the problem of **TRUTH BEARERS**: what are the bearers of truth? Some options:

- Propositions
- Sentences
- Utterances

But even if we agree that propositions are the bearers of truth, we can say that propositions are what sentences **EXPRESS**, and ask the question:

Does Tarski show to define predicates that have the same meaning as the English predicate 'expresses a truth'?

Another problem: 'true' (or 'expresses a truth') applies more widely than Tarski's predicates. It applies to the sentences of all languages — including English!

As we saw, Tarski thinks that this is the source of the Liar Paradox.

If he is right, his predicates *cannot* be co-extensive with 'true' (or 'expresses a truth').

## EXPLICATING TRUTH

Did Tarski show how to *explicate* truth, i.e. define predicates that can replace 'true' (or 'expresses a truth') in all legitimate theoretical contexts, but that lack its defects?

Here are four problems for this idea...

First, a problem raised by Michael Dummett: the **PROBLEM of TRUTH CONDITIONAL SEMANTICS**.

Many philosophers and linguists think that the notion of truth has an important theoretical role to play in an account of what the sentences of a language *mean*: "to know what a sentence means is to know its truth conditions".

But if a theory that tells us what the truth conditions of the sentences of a language is to serve as an account of what they mean, we have to *already* know the point of describing sentences as true.

Compare: being told their peng conditions.

So a Tarskian truth definition cannot *both* tell us what the sentences of the object language mean *and* tell us the meaning of the predicate being defined.

A possible response: reject the assumption that to know the meaning of a sentence is to know its truth condition.

Second, a problem raised by Hilary Putnam: the **MODAL PROBLEM**.

Remember our definition of 'true<sub>L<sub>1</sub></sub>':

$\forall s$  (s is true<sub>L<sub>1</sub></sub> IFF: either

- s = '1 + 1 = 2' and one plus one is two OR
- s = '1 + 1 = 3' and one plus one is three)

Let **TWO** be a name of '1 + 1 = 2'. Then, replacing 'true<sub>L<sub>1</sub></sub>' in 'TWO is true<sub>L<sub>1</sub></sub>' with its *definiens*, we get:  
'TWO = TWO and one plus one is two or TWO = '1 + 1 = 3' and one plus one is three'

The problem is that the latter is true in all possible worlds in which one plus one is two (i.e. in all possible worlds.)  
But isn't **TWO** *false* in some possible worlds — e.g. certain worlds in which '2' refers to the number three?

A possible response: deny that **TWO** is false in some possible worlds by insisting that sentences are individuated by their semantics as well as their syntax.

Third, a problem raised by Max Black: the **PROBLEM OF NON-PROJECTABILITY**.

The definitions that Tarski gives us don't *project*. That is, they don't tell us under what conditions other truth predicates, for other languages, hold of the sentences of those languages.

Fourth, a problem stressed by Scott Soames: the **PROBLEM OF EPISTEMIC DIFFERENCE**.

Knowing the conditions under which a sentence is true, whether or not it is sufficient for knowing what it means, at the very least provides information about what it does *not* mean.

For example, if one knows that 'la neige est blanche' is true IFF snow is white, one is at least in a position to know that it does *not* mean that snow is not white.

But knowing the conditions under which a sentence belongs to the extension of a Tarskian truth predicate does not provide even this sort of negative information about its meaning.

Suppose I tell you that 'yksi plus yksi on kaksi' is totta IFF one plus one is two. Are you able to determine whether 'yksi plus yksi on kaksi' means that one plus one is not two?

It seems not. Even if I give you a Tarski-style definition of 'totta', this doesn't rule out the possibility that 'yksi plus yksi on kaksi' means that one plus one is not two.

## THE WATERGATE EXAMPLES

Tarski *does* give us predicates that apply to all and only the true sentences of various languages without giving rise to the Liar Paradox. But his approach is very restrictive.

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

- 'true<sub>0</sub>' (roughly: 'is a true sentence not containing any truth predicate')
- 'true<sub>1</sub>' ('is a true sentence whose only truth predicates are "true<sub>0</sub>"')
- 'true<sub>2</sub>' ('is a true sentence whose only truth predicates are "true<sub>0</sub>" or "true<sub>1</sub>"')
- and so on...

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

Kripke's own theory promises to do better. Let's take a look.

## KRIPKE'S THEORY

## GROUNDEDNESS

One way to avoid the Liar Paradox is by denying, or at least refusing to accept, that the liar sentence is either true or false. (We'll see next week how exactly this helps.)

But if this to be anything more than an *ad hoc* response, we need some principled reason for thinking that it is correct.

Can we provide one?

There is an intuitive sense in which the truth of some but not all sentences depends on the truth of *other* sentences.

Compare:

- 'Snow is white'
- "'Snow is white' is true'

Whether "'Snow is white' is true' is true depends on whether 'Snow is white' is true. But the truth of 'Snow is white' doesn't depend on the truth of another sentence in the same way. It depends on whether snow is white.

These sentences form the initial segment of a certain hierarchy of sentences (the **SNOW HIERARCHY**) each successive member of which depends for its truth, in the relevant sense, on whether the previous member was true:

- 'Snow is white'
- "'Snow is white' is true'
- ""'Snow is white' is true' is true'
- ...

Since the truth of the first member does not depend on the truth of any other sentence, there is reasonably clear sense in which each member of the hierarchy is **GROUND**ED.

This suggests: the principled reason for refusing to accept that the liar sentence is true or false is that it is **UNGROUND**ED. Kripke spells this out more precisely.

## THE MINIMUM FIXED POINT

First, we construct a definition of 'true<sub>L</sub>' that applies to those sentences of L that do not contain the word 'true<sub>L</sub>'.

These are assigned to either the **EXTENSION** of 'true<sub>L</sub>', the set of sentences to which it applies, or to its **ANTI-EXTENSION**, the set of sentences to which it does not apply.

We then extend the assignment so that it covers sentences that **DO** contain the word 'true<sub>L</sub>', treating sentences that not (yet) belong to either its extension or its anti-extension as neither true nor false.

For this extension, we need a logic in which sentences can be neither true nor false. Kripke uses the Kleene strong logic, in which sentences are true, false, or indeterminate.

Thus, at this stage:

- 'Snow is white' is true
- "'Snow is white' is true<sub>L</sub>' is indeterminate
- In K<sub>3</sub>, 'Snow is white or 'Snow is white' is true<sub>L</sub>' is therefore true.

(The choice of K<sub>3</sub> here is not essential — any logic that allows sentences to be in some sense neither true nor false will do. Kripke also considers a **SUPERVALUATIONIST** approach.)

This gives us a level 0 of sentences: those whose truth does not depend on the truth values of *other* sentences.

In other words: the sentences that are assigned to either the extension or the anti-extension of 'true<sub>L</sub>' at this initial stage.

The English sentences at level 0 will include:

- 'Snow is white'
- 'Grass is red'
- 'Snow is white or most of Nixon's Watergate-related statements are not true'

But not:

- "'Snow is white' is true"
- 'Most of Nixon's Watergate-related statements are not true'
- 'Grass is red or most of Nixon's Watergate-related statements are not true'

Notice that whether or not a sentence belongs to level 0 is not determined by its syntax, and in particular by whether or not it contains the word 'true'!

The process can then be repeated on through subsequent levels:

Level 1: sentences whose truth values are determined by those of the sentences at level 0. For example:

- "'Snow is white' is true"
- "'Snow is white' is true or most of Nixon's Watergate-related statements are not true"

Level 2: sentences that do not belong to lower levels whose truth values are determined by the truth values of sentences at those lower levels. For example:

- "'Snow is white' is true' is true"
- "'Snow is white' is true' is true or most of Nixon's Watergate-related statements are not true"

The process goes on beyond all finite levels, accommodating examples like:

- 'Every sentence in the snow hierarchy is true'

Eventually, Kripke shows, we will reach a **MINIMUM FIXED POINT** at which no new sentences get added to the extension and anti-extension of 'true<sub>L</sub>'.



The liar sentence turns out to be ungrounded in the precise sense that it is not assigned to either the extension or anti-extension of 'true<sub>L</sub>' at any level in this process.

Neither are non-paradoxical but otherwise defective sentences such as:

- 'This sentence is true'
- 'This sentence is either true or false'

On the other hand, if most of Dean's Watergate-related statements are true, and most of Nixon's are false, both of the following will be found in the extension of 'true':

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

Since the defectiveness of liar sentences is not treated as syntactic, this is a far less restrictive approach.