Paradoxes

Lecture Eight

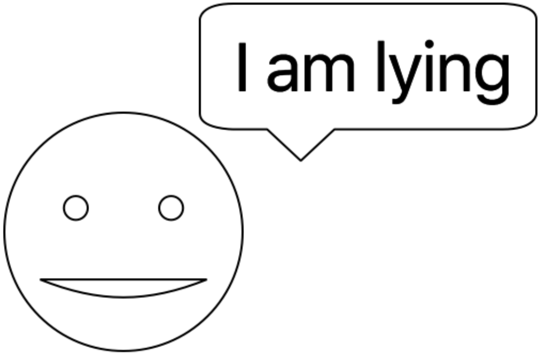
The Liar Paradox

Rob Trueman
rob.trueman@york.ac.uk

University of York

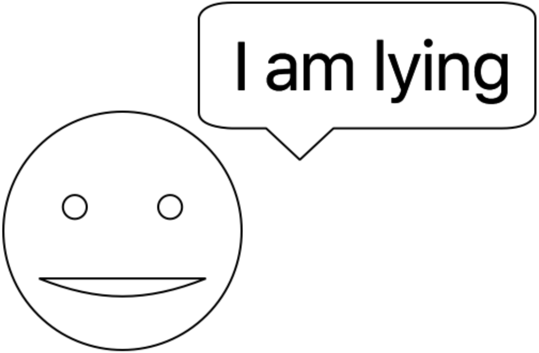
The Liar Paradox

The Liar Paradox



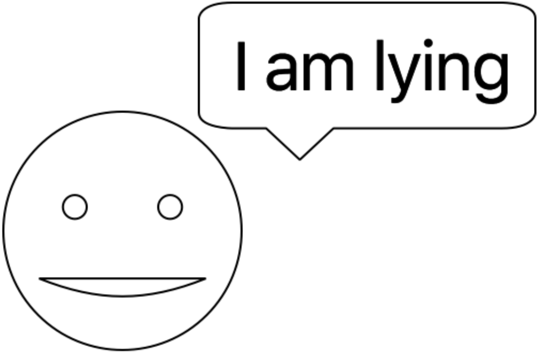
- A man says that he is lying. Is what he says true or false?

The Liar Paradox



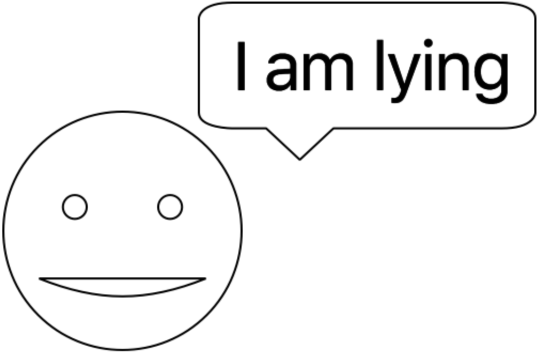
- If it is true, then he is lying; so it is false

The Liar Paradox



- If it is a lie, then he is not lying; so he is telling the truth

The Liar Paradox



- The man is lying iff he is telling the truth

A Pointless Riddle?

> *Why do you bore me with that which you yourself call the "liar fallacy"? [...] Not to know [such paradoxes] does no harm, and mastering them does no good.*
>
> *(Moral letters to Lucilius by Seneca, Letter 45: On sophistical argumentation)*

- No one agrees with Seneca anymore
- In the early 20th Century, there were **huge** advances in logic, and many of them began with reflections on the Liar Paradox
- However, if we are going to get anywhere, then first we need to get the paradox in its neatest, sharpest form

From Lying to Falsehoods

- Talk of ‘lying’ is an unnecessary complication, since lying involves **intention**
  - You are not lying if you say something false but think you are saying something true!

- We can formulate the Liar Paradox without getting involved with this complication

  What I am saying now is false

  (λ) λ is false

  The 6th sentence on this slide is false

- For now we will work with λ, but we must always remember that there are these other forms of the paradox

The Liar

- $\lambda =$ 'λ is false'

(T) '*p*' is true if and only if *p*

  - 'Snow is white' is true if and only if snow is white
  - 'The Paradoxes lectures have been excellent' is true if and only if the Paradoxes lectures have been excellent

- 'λ is false' is true if and only if λ is false

- λ is true if and only if λ is false

Truth-Value Gaps?

- λ is true if and only if λ is false
- This is paradoxical if we assume that λ must either be true or false
  - Suppose that λ is true; in that case λ is false; Contradiction!
  - Suppose that λ is false; in that case λ is true; Contradiction!
  - So either way, if λ is true or λ is false, then we are led to a contradiction!
- But if we say that λ is **neither** true **nor** false, then we cannot derive a contradiction
- Does this give us a quick response to the Liar Paradox?
  - λ is a truth-value **gap**, it is neither true nor false

The Strengthened Liar

- $\lambda =$ 'λ is not true'

(T) '*p*' is true if and only if *p*

- 'λ is not true' is true if and only if λ is not true
- λ is true if and only if λ is not true

Revenge Paradox

- We cannot stop this version of the paradox just by saying that λ is neither true nor false
  - If λ is neither true nor false, then it is not true
  - But that is exactly what λ says!
  - So if λ is neither true nor false, then it is true!
- This is known as a **revenge** paradox, and they come up all the time when we're dealing with the Liar

The Liar Paradox

Quinean Classifications

A paradox is an apparently unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises. (Sainsbury's definition from Paradoxes, p. 1)

- Premise-flawed
  - One of the premises turns out to be false
- Fallacious ("Falsidical")
  - The reasoning turns out to be faulty
- Veridical
  - The conclusion turns out to be true

Could the Liar Paradox be Veridical?

- Could the Liar Paradox be veridical?
- Surely not: the conclusion is a contradiction!
- Actually, the Liar Paradox has proven so tricky to deal with, that some philosophers (most notably Graham Priest) have been driven to say that some contradictions can be true
- This is a **very** radical response to the paradox, and we'll come back to it in Lecture 9

The Premises of the Liar Paradox

- So we must say that the paradox is either premise-flawed, or fallacious
- In this lecture, I am going to focus on solutions which say that it is premise-flawed
- The Liar Paradox has just two premises:
  (1) $\lambda =$ 'λ is not true'
  (2) 'λ is not true' is true iff λ is not true
- Which of these premises are we going to reject?

Is Self-Reference to Blame?

(1) $\lambda =$ 'λ is not true'

- (1) captures a very weird fact about λ
- λ is **self-referential**: it talks about itself
  - λ says that λ is not true
- We might have thought that the Liar Paradox is just a proof that this sort of self-reference is impossible

Self-Reference and Arithmetic

- There are ways of ‘coding up’ claims about language into arithmetic
- This took a genius (Gödel) to realise, but now the idea is very familiar
  - Think of the way that computers encode information into 1s and 0s
- When we code things up in this way, we find that arithmetic allows us to create a version of λ, which is self-referential in just the way that is needed for the Liar Paradox
- So unless we want to challenge ordinary arithmetic, we cannot complain about the kind of self-reference used in the Liar Paradox

In Defence of Self-Reference

- Also, remember that we have to deal with paradoxes like this:
  (a) The 2nd sentence on this slide is not true
- (a) is a self-referential liar paradox, but only contingently; to see this, compare it with (b):
  (b) The 2nd sentence on the previous slide is not true
- (b) is a perfectly meaningful, intelligible sentence
- We have no difficulty assigning it a truth-value (it is false)
- But if we are allowed to use sentences like (b), then we cannot eliminate contingent liar sentences like (a)

The Premises of the Liar Paradox

(1) ~~$\lambda =$ ‘λ is not true’~~

(2) ‘λ is not true’ is true iff λ is not true

- So if we are going to reject any premise, it will have to be (2)
- (2) is an instance of the general schema (T):

  (T) ‘p’ is true iff p

- So we need to take a closer look at (T)

The Liar Paradox

Semantically Closed Languages

- According to Tarski, the Liar Paradox comes up because we are trying to use English to talk about truth-in-English
- In Tarski's terminology, the Liar Paradox comes up (in English) because English is **semantically closed**
- A language is **semantically closed** iff
  - (i) it contains its own truth predicate satisfying (T); and
  - (ii) it contains the means to form names for its own sentences (e.g. via quotation)

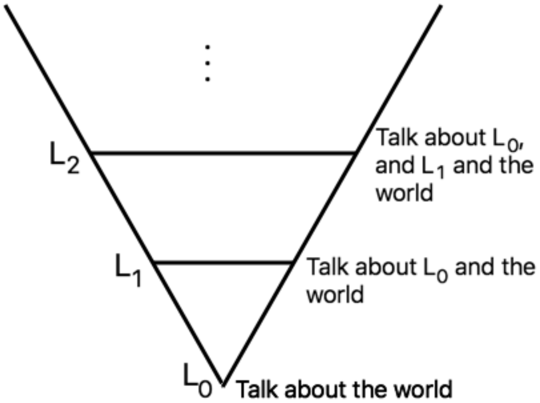
Avoiding Semantic Closure

- For Tarski, the lesson of the Liar Paradox was that natural languages like English, which are semantically closed, are irredeemably inconsistent
- Tarski abandoned natural languages, and instead studied how we could create consistent **formal** languages
- His idea was that there is not one over-arching notion of truth, but a different notion of truth for each formal language L, **truth-in-L**

Avoiding Semantic Closure

- No language is ever allowed to contain its own truth predicate
- If you want to talk about truth-in-L_0, then you need to move up into a more powerful language, L_1
- We can use L_1 to talk about truth-in-L_0, but if we then want to talk about truth-in-L_1, then we need to move up to an even more powerful language, L_2, and so on

A Hierarchy of Languages



Back to the T-Schema

(T) '*p*' is true iff *p*

- We should swap (T) for this schema:

(T_n) '*p*' is true-in-L_n iff *p*

- Since this talks about truth-in-L_n, the sentence '*p*' must be from L_n
- But (T_n) itself belongs to L_{n+1}, since it mentions truth-in-L_n
  - I am simplifying a bit by writing (T_n) like this
  - I am assuming that L_n is an **extension** of L_{n-1}, so that L_n contains all the expressions of L_{n-1} (and then some)
  - This does not have to be the case, but it makes things a lot simpler

Some Examples

(T_n) 'p' is true-in-L_n iff p

- 'Snow is white' is true-in-L_0 iff snow is white
  - This is an acceptable instance of (T_n), because 'Snow is white' is a sentence from L_0
- ' "Snow is white" is true-in-L_0' is true-in-L_1 iff 'Snow is white' is true-in-L_0
  - This is also an acceptable instance of (T_n), because ' "Snow is white" is true-in-L_0' is a sentence in L_1
- ' "Snow is white" is true-in-L_0' is true-in-L_0 iff 'Snow is white' is true-in-L_0
  - This is an **unacceptable** instance of (T_n), because ' "Snow is white" is true-in-L_0' is **not** a sentence in L_0

Solving the Lair Paradox

- How does this help with the Liar Paradox?
- The Strengthened Liar was meant to say of itself that it was not true:
  - $\lambda =$ 'λ is not true'
- But now we need to replace 'true' with 'true-in-L_n':
  - $\lambda_n =$ 'λ_n is not true-in-L_n'

Solving the Lair Paradox

- To get a paradox going, we would then need to take the following instance of (T_n):
  - 'λ_n is not true-in-L_n' is true-in-L_n if and only if λ_n is not true-in-L_n
- But this is an **unacceptable** instance of (T_n)
- λ_n cannot be a sentence of L_n, because it mentions truth-in-L_n!

The Liar Paradox

Overkill?

- Everyone accepts that Tarski's solution to the Liar Paradox works in the sense that it gets us out of the paradox
- But lots of philosophers also think that the solution is too extreme
- In effect, Tarski blocks the paradox by restricting our expressive powers
  - We can no longer talk about truth full-stop, only truth-in-L; and we can never talk about truth-in-L in L
- This prevents us from saying lots of things that we want to be able to say

Bivalence

- Sometimes we want to make absolutely general claims about truth, like:
  (1) Every sentence is either true or false
- This claim (**The Principle of Bivalence**) may be true or it may be false, but it is certainly a claim that we want to be able to discuss
- But on Tarski's picture, we can't; all we *can* say is:
  (2) Every sentence in L_n is either true-in-L_n or false-in-L_n
- Moreover, (2) will belong to L_{n+1}, which is a language that (2) cannot be applied to

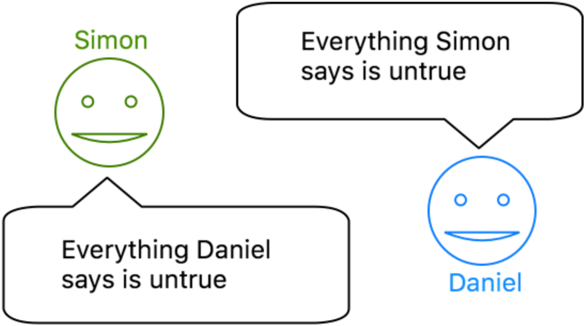
Applying Tarski to Natural Languages

- Tarski explicitly set natural languages to one side; his theory of truth was meant to deal with formal languages
- But what are we to say about natural languages?
- Some philosophers have tried to re-apply what Tarski said about formal languages
- The idea is that English itself is a Tarskian hierarchy of languages

Applying Tarski to Natural Languages

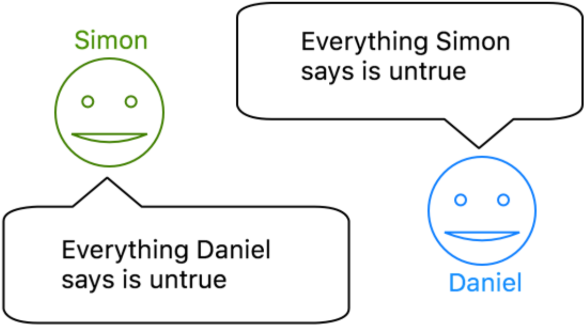
- We have a fragment of English which talks just about the world, and not truth, L_0
- Then we have a fragment of English which talks about the world and truth-in-L_0, L_1
- Then we have a fragment of English which talks about the world and truth-in-L_0 and truth-in-L_1, L_2
- ...
- The English word 'true' is ambiguous between 'true-in-L_0', 'true-in-L_1', and so on

A Problem for Applying Tarski to Natural Languages



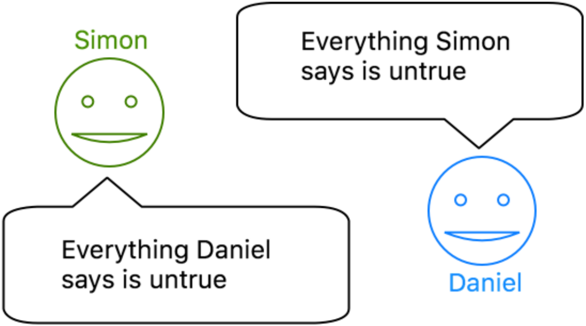
- Simon intends his assertion to apply to **everything** that Daniel says

A Problem for Applying Tarski to Natural Languages



- Daniel intends his assertion to apply to **everything** that Simon says

A Problem for Applying Tarski to Natural Languages



- But there is no consistent way of assigning levels which has this effect

A Problem for applying Tarski to Natural Languages

(1) Everything Daniel says is untrue

(2) Everything Simon says is untrue

- Suppose that Simon means true-in-L_0 by 'true'
- In that case, (1) must belong to L_1
- If Daniel wants his assertion to apply to (1), he must have meant untrue-in-L_1 by 'untrue'
- So (2) must belong to L_2
- Thus (1) does not apply to (2)!

The Liar Paradox

The Strengthened Liar

- The most difficult version of the Liar Paradox starts with the **Strengthened Liar**:

  (λ) λ is not true

- λ involves a peculiar kind of self-reference, and it is tempting to block the Liar Paradox by outlawing this kind of self-reference

- Unfortunately, Gödel showed us that we can code up claims about language into numbers, and that such a coding allows us to cook up self-referential sentences

- And even if we didn't like any of that Gödel stuff, we could cook up Contingent Liar Paradoxes
  - The 6th sentence on the this slide is not true

A Balancing Act

- The only other option is to somehow restrict schema (T)

  (T) '*p*' is true iff *p*

- Finding the right restriction is a balancing act
  - The harsher the restriction on (T), the more limited our truth-talk will be
  - So we need to find a restriction on (T) which blocks the paradox, but which doesn't limit our truth-talk too badly

- Tarski's restriction on (T) blocks the paradox, but the counter-examples suggest that it may be *too restrictive*

Where Next?

- We could try to learn to live within Tarski's fairly restrictive linguistic heirarchy
- Or we could find some other way of blocking the Liar Paradox which does not impose such harsh restrictions
  - See the **bonus slides** on Kripke's theory, which he developed in 'Outline of a Theory of Truth'
- Or, most radically, we could try to escape this balancing act altogether by just accepting a contradiction: λ is true **and** not true
  - This is Priest's response to the paradox, and we will look at it next week

This Week's Seminar

- **Required Reading:**
  - *Paradoxes* ch. 6, §§6.2–6.8
  - J.L. Mackie, *Truth, Probability and Paradox*, ch. 6, §§1–5
- Please think through the questions outlined in the last VLE announcement
- Also write **at least** one question you would like answered about Russell's Paradox or the Liar Paradox, and bring it to the seminar

Next Week

- Next week we will be consider a radical solution to paradoxes like the Liar and Russell's Paradox: accepting the existence of true contradictions!
- Required reading:
  - *Paradoxes* chapter 7

References

- Kripke, S (1975) 'Outline of a Theory of Truth', *Journal of Philosophy* 72: 690–715
- Tarski, A 'The concept of truth in formalized languages' in Corcoran (ed) Woodgar (trans) *Logic, Semantics, Metamathematics* (Indianapolis, IN: Hackett, 1983) pp. 152–278