How to Define Truth-in-L

(Tarski's semantic definition of truth simplified)

April 3, 2007

Let L be the language for which you're trying to define truth: the object language. Let M be the language in which you're framing the definition: the metalanguage. (Tarski showed that in general, M cannot be L. But M can include L.)

Tarski's goal: to define a predicate 'true-in-L' that is true of every true sentence of L and false of everything that is not a true sentence of L.

Note: sentences (as we understand them here) are just grammatical sequences of words. We name them by putting them in quotation marks: thus, for example, 'Snow is white' is the name of the sentence you get by concatenating the words 'Snow', 'is', and 'white' in that order.

The adequacy question: How do you know when you've defined 'true-in-L' correctly?

Convention T. A definition of 'true-in-L' is correct only if it implies every sentence of the form

(*) S is true-in-L iff p,

where S is the name of a sentence in L and p is a translation of that sentence into M.

Sentences of the form (*) are called **T-sentences** or **T-biconditionals**. Some examples:

- 'Snow is white' is true-in-English iff snow is white.
- 'Schnee ist weiss' is true-in-German iff snow is white.
- 'Only cowards eat cupcakes after swimming' is true-in-English iff only cowards eat cupcakes after swimming.

You can think of the T-sentences as the "data" against which a definition of 'true-in-L' is to be tested.

Why this is nontrivial. You might think that it's easy to give a definition of 'true-in-L' that is adequate by the lights of Convention T. Just give the definition in a metalanguage M that includes L (i.e., every sentence of L is also a sentence of M, with the same meaning). Then define as follows:

(E) for all sentences S in L, S is true-in-L iff S.

Doesn't (E) meet the criterion of adequacy set down by Convention T?

No. (E) is actually ungrammatical. The variable 'S' on the left-hand side of the biconditional must refer to a sentence in L. But 'S' on the right- hand side of the biconditional must be a sentence in L.

Somehow, the definition has to match a sentence of M with each name of a sentence in L. There are infinitely many sentences in L, so we can't just list these correspondences. Somehow, we have to specify them in a finite way.

Tarski's approach is to define 'true-in-L' *recursively*. That is, he gives you a procedure for figuring out the T-biconditional for a complex sentence if you already know the T-biconditionals for simpler sentences out of which it is formed. (This is not quite right—I'm leaving out some complexity, which has to do with Tarski's treatment of quantifiers like 'everything'—but it's close enough for our purposes.)

An example of a recursive definition with which you're all familiar:

> (R1) $y^0 = 1$
> (R2) $y^{n+1} = y^n \times y$

Using only two sentences, we've specified the value of y^n for any n. We've done this by giving a recipe for figuring out y^{n+1} if we already know y^n. Suppose we want to figure out y^3. (R2) tells us that $y^3 = y^2 \times y$. So now all we need to figure out is y^2. (R2) tells us that $y^2 = y^1 \times y$. So we know that $y^3 = (y^1 \times y) \times y$, and now all we need to figure out is y^1. (R2) tells us that $y^1 = y^0 \times y$, so we know that $y^3 = ((y^0 \times y) \times y) \times y$. And (R1) tells us that $y^0 = 1$, so $y^3 = (((1 \times y) \times y) \times y) \times y = y \times y \times y$. Tedious, but it works!

A sample recursive truth definition: Let L be a language with

- two nouns, 'owsnay' meaning *snow*, and 'oalcay' meaning *coal*
- two predicates, 'isyay iteway' meaning *is white*, and 'isyay ackblay' meaning *is black*
- one unary connective, 'otnay' meaning *it is not the case that*
- one binary connective, 'andyay' meaning *and*

Grammar of L: A sentence of L is...

- a noun concatenated with a predicate, or
- a unary connective concatenated with a sentence, or
- a binary connective concatenated between two sentences.

(Note that there are four simple noun-predicate sentences of L but infinitely many sentences altogether.)

Let M (our metalanguage, in which the definition will be stated) be English.

Recursive truth definition for L:

1. 'owsnay isyay itewhay' is true-in-L iff snow is white.
2. 'owsnay isyay ackblay' is true-in-L iff snow is black.
3. 'oalcay isyay itewhay' is true-in-L iff coal is white.
4. 'oalcay isyay ackblay' is true-in-L iff coal is black.
5. If A is a sentence, then 'otnay' concatenated with A is true-in-L iff A is not true-in-L.

6. If A, B are sentences, then A concatenated with ‘andyay’ concatenated with B is true-in-L iff A is true-in-L and B is true-in-L.

Let’s try it. Take a random sentence of L, say, ‘Owsnay isyay itewhay andyay otnay oalcay isyay itewhay’.

- By clause 6, ‘Owsnay isyay itewhay andyay otnay oalcay isyay itewhay’ is true-in-L iff ‘Owsnay isyay itewhay’ is true-in-L and ‘otnay oalcay isyay itewhay’ is true-in-L.
- By clause 5, ‘otnay oalcay isyay itewhay’ is true-in-L iff ‘oalcay isyay itewhay’ is not true-in-L.
- So, putting together the last two steps, ‘Owsnay isyay itewhay andyay otnay oalcay isyay itewhay’ is true-in-L iff ‘Owsnay isyay itewhay’ is true- in-L and ‘oalcay isyay itewhay’ is not true-in-L.
- By clause 1, ‘Owsnay isyay itewhay’ is true-in-L iff snow is white.
- By clause 3, ‘oalcay isyay itewhay’ is not true-in-L iff coal is not white.
- So, putting together the last three steps, ‘Owsnay isyay itewhay andyay otnay oalcay isyay itewhay’ is true-in-L iff snow is white and coal is not white.

Note that this is a T-biconditional, and the right-hand side is a translation into M of the sentence of L named in the left-hand side. You can convince yourself that the truth definition can be used in this way to derive all the other such T-biconditionals for L. So it’s adequate by the lights of Convention T.

This may seem trivial—and admittedly L isn’t much of a language—but look how much we’ve done. We’ve given a precise definition of ‘true-in-L’ without helping ourselves to any semantic language at all (truth, reference, etc.). This finite definition tells us the condition under which any one of the infinitely many sentences of L is true.