

Tarski Undefinability Theorem Reexamined

Tarski proved that the Liar Paradox: $G \leftrightarrow \sim(F \vdash G)$ is true in his meta-theory and not provable in his theory without ever realizing that the only reason it is not provable in his theory is that it is not true in his theory.

Examining his Undefinability Theorem with the (Curry 2010) notion of a formal system: A theory T is a conceptual class consisting of certain of these elementary statements. The elementary statements which belong to T are called the elementary theorems of T and said to be true. In this way, a theory is a way of designating a subset of E which consists entirely of true statements.

This general way of designating a theory stipulates that the truth of any of its elementary statements is not known without reference to T. Thus the same elementary statement may be true with respect to one theory, and not true with respect to another. (Curry, Haskell. 2010 Foundations of Mathematical Logic).

We derive these three universal Truth predicate axioms:

- (1) $\forall F \forall x (\text{True}(F, x) \leftrightarrow (F \vdash x))$
- (2) $\forall F \forall x (\text{False}(F, x) \leftrightarrow (F \vdash \sim x))$
- (3) $\forall F \forall x (\sim \text{True}(F, x) \leftrightarrow \sim(F \vdash x))$

Tarski rejected the above truth predicates could possibly exist essentially on the basis of his assumption that they would prove the following sentence true: $\exists F \exists G (G \leftrightarrow \sim(F \vdash G))$. (Tarski page 248)

$G \leftrightarrow \sim(F \vdash G)$ Means that G has the same Truth value as its own unprovability in F. When the RHS $\sim(F \vdash G)$ is true, by Truth axiom(3) we know that x is not true in F. This contradicts the LHS being true, making the expression false.

Tarski then went on to prove the above sentence to be true in his meta-theory yet unprovable in the theory, never realizing that the only reason it is unprovable in the theory is that it is untrue in the theory (Tarski page 276).

Copyright 2019 Pete Olcott

Excerpts from “The concept of truth in formalized languages” Tarski 1936

// Tarski page 248

Should we succeed in constructing in the metalanguage a correct definition of truth, then ... It would then be possible to reconstruct the antinomy of the liar in the metalanguage, by forming in the language itself a sentence x such that the sentence of the metalanguage which is correlated with x asserts that x is not a true sentence.

// Tarski page 276

The formulas (8) and (9) together express the fact that x is an undecidable sentence; moreover from (7) it follows that x is a true sentence.

By establishing the truth of the sentence x we have eo ipso -by reason of (2)-also proved x itself in the metatheory. Since, moreover, the metatheory can be interpreted in the theory enriched by variables of higher order (cf. p. 184) and since in this interpretation the sentence x , which contains no specific term of the metatheory, is its own correlate, the proof of the sentence x given in the metatheory can automatically be carried over into the theory itself: the sentence x which is undecidable in the original theory becomes a decidable sentence in the enriched theory.

Proof on pages 275-276, x defined on page 248

http://www.thatmarcusfamily.org/philosophy/Course_Websites/Readings/Tarski%20-%20The%20Concept%20of%20Truth%20in%20Formalized%20Languages.pdf