Two Ways to Negate Self-Referential Sentences

David Wyde

September 29, 2020

Abstract

The negation of a self-referential sentence should refer to itself rather than to the original statement. This approach bypasses Gödel's incompleteness theorems and the liar paradox.

1 Background

The negation of a sentence states its opposite[1]. The negation of "5 is a prime number" is "5 is not a prime number."

If neither a sentence nor its negation can be proven, that sentence is undecidable[2]. An undecidable sentence is a glitch in a formal system, an input that the system cannot handle.

2 Gödel's Incompleteness Theorems

In a 1931 paper, Gödel proposed an undecidable sentence and used it as the basis of his two incompleteness theorems: 1) any formal system with certain powers is unable to prove some true propositions and 2) such a system cannot prove its own consistency (lack of contradictions).

In essence, Gödel's undecidable sentence is

Sentence 1: Sentence 1 is not provable.

Gödel used a particular scheme to number his statements[3]. This led to a sentence that asserts its own unprovability.

Gödel's idea is that Sentence 1 is undecidable because both it and its negation are unprovable. If we try to prove Sentence 1, it asserts something false (its inability to be proven). On the other hand, if we cannot prove Sentence 1, it is true but unprovable.

3 The Liar Paradox

Near the beginning of his 1931 paper, Gödel referred to the liar paradox[4]:

Sentence 2: This sentence is false.

Trying to evaluate Sentence 2 leads to a type of infinite loop: if it is true it must be false, and if it is false it must be true.

An important issue for both the liar paradox and Gödel's incompleteness theorems is how to negate Sentences 1 and 2. A key question is whether the negation of a self-referential sentence should refer to the initial sentence or to the negation itself.

We will start with the liar paradox, Sentence 2. Two possible negations are

Sentence 3: This sentence is true.

and

Sentence 4: Sentence 2 is true.

The self-referential Sentence 3 is true and provable, so it avoids a paradox.

Sentence 4 is not true, since we saw above that Sentence 2 is not true. In this interpretation, both Sentence 2 and its negation are not true. This result conflicts with the definition of negation.

Here we end up with a paradox from the assumption that the negation of a self-referential sentence should refer to the original sentence rather than to the negation itself.

4 Negating Gödel's Sentence

Next we will analyze Gödel's sentence. Two potential negations for Sentence 1 are

Sentence 5: Sentence 5 is provable.

and

Sentence 6: Sentence 1 is provable.

The negation that refers to itself, Sentence 5, is provable. This means that Sentence 1 is decidable, which sidesteps Gödel's incompleteness theorems.

Sentence 6, the traditional way to negate Sentence 1, is analogous to the liar paradox's Sentence 4. Sentences 1 and 6 are both unprovable. This matching unprovability makes Gödel's sentence undecidable.

To further compare the differences between these two ways to negate self-referential sentences, we can negate Sentences 5 and 6. This produces two candidates for the negation of Sentence 1's negation.

Preserving self-reference, the negation of Sentence 5 is

Sentence 7: Sentence 7 is not provable.

Other than the numeric label, this is identical to Sentence 1.

With Gödel's method, the negation of Sentence 6 is

Sentence 8: Sentence 1 is not provable.

By Gödel's first incompleteness theorem, Sentence 1 is not provable. So, Sentence 8 is both true and provable.

With Gödel's approach, Sentence 1 is unprovable and self-referential, while the negation of its negation is provable and not self-referential.

5 Conclusion

Gödel's undecidable sentence, used to prove fundamental limits on the power of logic, might be the result of choosing the wrong negation. The same analysis applies to the liar paradox.

References

[1] P. Suppes, *Introduction to Logic*. New York: Van Nostrand Reinhold Company, 1957, pp. 3–4.

[2] K. Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I," *Monatshefte für Mathematik und Physik*, vol. 38, pp. 173–198, 1931 (Transl.: B. Meltzer, *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. New York: Basic Books, 1962. Reprinted Mineola: Dover, 1992).

[3] E. Nagel and J. R. Newman, *Gödel's Proof*. New York: New York University Press, 2001, ch. VII.

[4] D. R. Hofstadter, *Metamagical Themas*. New York: Basic Books, Inc., 1985, pp. 7.