Summary of Godel's Incompleteness Theorem

sajal harmukh

22 July 2021

Godel's Incompleteness Theorems

Gödel's two incompleteness theorems are among the most important results in modern logic, and have deep implications for various issues. They concern the limits of provability in formal axiomatic theories. The first incompleteness theorem states that in any consistent formal system F within which a certain amount of arithmetic can be carried out, there are statements of the language of F which can neither be proved nor disproved in F.

1 Description of two theorems

The first incompleteness theorem states that no consistent system of axioms whose theorems can be listed by an effective procedure (i.e., an algorithm) is capable of proving all truths about the arithmetic of natural numbers. For any such consistent formal system, there will always be statements about natural numbers that are true, but that are unprovable within the system. The second incompleteness theorem, an extension of the first, shows that the system cannot demonstrate its own consistency.

2 Implication on Artificial Intelligence

ne of the goals of AI research is to achieve "strong artificial intelligence", meaning human-level general AI. Currently, we build AI as algorithms in Turing machines, which are consistent axiomatic systems and therefore subject to the theorem.

Roger Penrose and J.R. Lucas argue that human consciousness transcends Turing machines because human minds, through introspection, can recognize their own inconsistencies, which under Gödel's theorem is impossible for Turing machines. They argue that this makes it impossible for Turing machines to reproduce traits of human minds, such as mathematical insight.