Historia Mathematica Vol. 13, No. 1, 2000, 137-141

Gödel's Disjunctive Conclusion

연세대학교 인지과학연구소 **현우식**

Abstract

This paper discusses *Gödel's Disjunctive Conclusion* in terms of cognitive science and his Incompleteness Theorems from a metamathematical perspective.

0. Introduction

According to Gödel, there exist two alternatives on the equivalence of the mind and the machine. In 1951, Gödel delivered the 25th Josiah Willard Gibbs Lecture, entitled *Some Basic Theorems on the Foundations of Mathematics and Their Implications* [5], at the American Mathematical Society. In this talk, he addressed the significance of the incompleteness theorems for the controversies surrounding the nature of mathematics and the limitations of human cognition. He focused on the fundamental issue of the human mind and mechanical objects. Gödel asserted that the following disjunctive conclusion is inevitable with respect to the undecidable:

> *Either mathematics is incompletable in this sense, that its evident axioms can never be comprised in a finite rule, that is to say, the human mind (even within the realm of pure mathematics) infinitely surpasses the powers of any finite machine, or else there exist absolutely unsolvable diophantine problems of the type specified* (where the case that both terms of the disjunction are true is not excluded, so that there are, strictly speaking, three alternatives). It is this mathematically established fact which seems to me of great philosophical interest(Italics in original, [5, p. 310]).

For Gödel, it is *intuitionists* in the foundation of mathematics who assert the first alternative of the disjunction and negate the second part. Gödel regarded *finitists* as opponents of the first disjunctive term. Epistemologically, these terms are more precise expressions rather than mechanism or anti-mechanism. Thus for the intuitionist or finitist, the theorem holds as an implication instead of a disjunction [see [5], footnote 15].

Contrary to Penrose's claim [7], such a disjunctive conclusion illustrates that Gödel's position cannot be simply reduced to a mystical one.1) Rather, according to Gödel's interpretation, Penrose would be classified as an intuitionistic mathematician, and Turing a finitistic mathematician. Moreover, the disjunction may clarify some of the controversial terms in cognitive science, such as mechanism and anti-mechanism, or AI(Artificial Intelligence) thesis and anti-AI thesis.2)

Scholars like Hao Wang [2; 3], Rudy Rucker [8], and Stewart Shapiro [9] have concerned themselves with Gödel's disjunction and Gödel's view on human and artificial cognition. Wang [2] convincingly argued that Gödel affirmed neither *computabilism* nor *neuralism* as being true.3) We should note, however, that Gödel presupposed the brain to be a machine. Therefore, the term neuralism refers simply to Turing computabilism or a physical theory, rather than contemporary brain theories.

1. The First Alternative

The first alternative is that the mind is superior to the machine, whereas the second shows that there is no mind beyond the machine. Suppose that Gödel's second incompleteness theorem is true in both the human mind and any finite machine. Then there is a proposition decidable by the human mind, yet undecidable by any finite machine. A clear example of this is the proposition expressing the consistency of finite machines. In this sense, mathematics is *incompletable* because mathematics cannot be embodied in a finite rule. For Gödel, this consequently implies that the human mind cannot be reducible to any finite describable machine. Accordingly, there is no essential difference between the first alternative and Penrose's argument. Gödel remarked:

1) Penrose asserts that in a mystical viewpoint is Gödel's view concerning the relationship between mathematical thinking and computation: 'Awareness cannot be explained by physical, computational, or any other scientific terms.'[7]

2) Mechanism is the thesis asserting that the human mind and the brain are machines. The AI thesis[1, p.579] states: As the intelligence of machines evolves, its underlying mechanisms will gradually converge to the mechanism underlying human intelligence.

3) In his work [2], computabilism refers to the thesis that the brain and the mind function basically like a computing machine. Nueralism is the thesis that the brain suffices to explain mental phenomena.

> Namely, if the first alternative holds, this seems to imply that the working of the human mind cannot be reduced to the working of the brain, which to all appearances is a finite machine with a finite number of parts, namely, the neurons and their connections[5, p. 311].

This first alternative, however, does not necessarily mean that the incompleteness theorems preclude the existence of an idealized AI produced by a finite rule. Rather, the theorems say that if there is such a finite rule then we will be unable to recognize it as such. As Gödel stated,

> It is not known whether the first alternative holds, but at any rate it is in good agreement with the opinions of some of the leading men in brain and nerve physiology, who very decidedly deny the possibility of a purely mechanistic explanation of psychical and nervous process[5, p. 312].

Gödel never asserted that the first alternative holds in terms of his incompleteness theorems.

2. The Second Alternative

For the second alternative term, Gödel's theorem shows that there exists an undecidable proposition for both the human mind and any finite machine if *a specified type* is allowed; otherwise we cannot use Gödel's results. This is an essential point in the application of Gödel's theorems to the issues being discussed here. Therefore, there is no essential difference between this alternative and Turing's argument that mental procedures cannot go beyond mechanical procedures. Gödel asserted that the first alternative seems to imply that

> … mathematical objects and facts (or at least *something* in them) exist objectively and independently of our mental acts and decision(Italics in original, [5, p. 311]).

For Gödel, this case disproves the view that mathematics is merely our own creation. However, to verify this alternative, we must first specify the type of system.

In his work [6], Gödel pointed out that Turing's argument [10] - mental procedures cannot go beyond mechanical procedures - is inconclusive. Gödel claimed that

> Turing completely disregarded the fact that mind, in its use, is not static, but constantly developing, i.e., that we understand abstract terms more and more precisely as we go on

using them[6, p. 306].

Gödel, therefore, did not accept Turing's argument. From the viewpoint of the incompleteness theorems, however, Gödel never claimed that they refuted Turing's mechanistic view of the mind, which is a different assertion altogether. He asserted only that the disjunctive conclusion must hold.

3. The Third Alternative

Since both terms of the disjunction are true is not excluded, the third alternative follows. Gödel focused on the disjunction, whereas intuitionists and finitists focused on only one alternative term of the disjunction. They would regard Gödel's conclusion not as a disjunction but as an implication. Gödel, however, never claimed that the incompleteness theorems refute one alternative view of the mind. Rather, he refuted such an logical implication. According to Gödel [5], the disjunctive conclusion must hold:

> Either (1) the human mind infinitely surpasses the powers of any finite machine, or else (2) there exist absolutely unsolvable diophantine problems.

Gödel's own interpretations allowed a corresponding disjunction:

> Either (1) the working of the human mind cannot be reduced to the working of a finite machine, or else (2) mathematical objects and facts exist objectively and independently of our acts and decision.

Of course, those alternatives were not mutually exclusive. On this point, I agree with Dawson [4] that, indeed, Gödel was firmly convinced of the truth of both. Thus, follows a disjunction for cognitive science that

> Either (1) the working of human mind cannot be reduced to the working of a Turing machine, or else (2) Gödel sentences exist independently of our cognition.

4. Conclusion

In all, one can only say that Gödel maintained the disjunctive conclusion not an implication. We, therefore, cannot simply say that Gödel's incompleteness theorems imply only that the human mind is superior to the artificial mind. Gödel's disjunctive conclusion does not imply one specific alternative. Rather, it stands in contrast with a fixed alternative.

Gödel's own views on the mind, the brain, and the machine reveal new issues with respect to the incompleteness theorems. The problem essentially involves the finite, the specification of type and consistency. Before a decision can be made as to which alternative to follow, difficult questions regarding finitude and consistency must first be considered.

References

1. D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Basic Books Inc., New York: 1979.
2. H. Wang, *A Logical Journey: From Gödel to Philosophy*, The MIT Press, Cambridge, 1996.
3. H. Wang, *From Mathematics to Philosophy*, Routledge and Kegan Paul, London, 1974.
4. J. W. Dawson, Jr., *Logical Dilemmas: The Life and Work of Kurt Gödel*, A K Peters, Massachusetts, 1997.
5. K. Gödel, "Some Basic Theorems on the Foundations of Mathematics and Their Implications (1951)," *Collected Works volume III: Unpublished Essays and Lectures,* S. Feferman et al.(eds.), Oxford University Press, Oxford, 1995.
6. K. Gödel, "Some Remarks on the Undecidability Results (1972)," *Collected Works volume II: Publications 1938-1974,* S. Feferman et al.(eds.), Oxford University Press, Oxford, 1990.
7. R. Penrose, *Shadows of the Mind*, Oxford University Press, Oxford, 1994.
8. R. Rucker, *Infinity and the Mind: The Science and Philosophy of the Infinite*, Princeton University Press, New Jersey, 1995.
9. S. Shapiro, "Incompleteness, Mechanism, and Optimism," *The Bulletin of Symbolic Logic*, 4(3) (1998), 273-302.
10. A. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society*, 42 (1936), 230-265.