PHIL 261 Essay 1

The question that drives this paper is ‘what (if any) consequences do Gödel’s incompleteness results have for our understanding of the relationships between human minds, mathematics and machine computation’? Gödel’s incompleteness results developed in response to Hilbert’s five pillar program which sought to secure the foundation of mathematics while circumventing the introduction of any untenable paradoxes. Specifically, Gödel’s incompleteness results eradicated Hilbert’s firm foundation and revolutionized the philosophy of mathematics. Through his reasoning with Peano Arithmetic, Gödel’s Incompleteness Theorem provides greater ways to think about mathematics, artificial intelligence and the human mind. In this essay, I argue that his conclusions imply that mathematics requires an understanding of semantic content and that no formal system is able to encompass complete human reasoning. I further show why these factors limit the quest for general artificial intelligence and why the human mind is superior to the highest possibilities of machine cognition.

Gödel revolutionized the philosophy of mathematics when he succeeded in proving that there are truths in mathematics that lie beyond the realm of mathematical logic. Specifically, his Incompleteness Theorem showed that a purely formal system fails to encapsulate all the expressible truths within that system using its own rules and proofs[5]. Secondly, he showed that a mathematical system cannot prove its own consistency[5]. Additionally, his incompleteness results applies to systems of every scale including the human mind furthering the potency of its implications. One critical implication of his results is that truth and proof are separable[6]. This is illustrated in Penrose’s example of “G” whereby the statement “G” is true in our minds but not provable[6]. This shows how truth is accessible to humans while failing to be accessible to tools that rely strictly on logic, such as machines. If human minds were strictly computational then human minds would not have access to a truth statement that is ungrounded by a proof. Thus, the human mind must not be a strictly computational tool. This is further supported by Moore who highlights why mathematics requires marrying syntax (logical rules) with semantics (meaning of objects): “Because what Hilbert had envisaged - at least as a paradigm - was a single, complete axiomatization. And Gödel’s theorem shows that nothing matches that paradigm. Any axiomatic base for transfinite mathematics must be supplemented. Not only that, but there will be one particular way of supplementing it that seems forced upon us; and this casts doubt on the idea that only finitary propositions genuinely describe mathematical reality”[3]. Thus, the Gödel results demonstrate that syntax is insufficient to encapsulate mathematical reasoning, mathematics must have semantic content. This conclusion has led philosophy of mathematics towards the view of mathematical platonism, although this view will not be discussed further in this paper.

Beyond the mathematical implications, the Gödel results have significant implications on the field of artificial intelligence, namely, because no single formal system is able to capture the entirety of human reasoning. Under this view, the quest for general artificial intelligence is rendered futile. Although this is a limiting conclusion for the progression of artificial intelligence, it provides the opportunity to refocus research in artificial intelligence accordingly as per “this does not mean that a machine cannot simulate any piece of mind; it only says that there is no machine that can simulate every piece of mind”[1]. Gödel further illustrates the dilemma that his results pose to our understanding of the human mind in relation to machine cognition by stating: “So the following disjunctive conclusion is inevitable: 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 . . . "[2]. This implies either (1) the mind is not a machine, or (2) it is a machine and there are undecidable mathematical problems. Penrose's earlier example of "G" illustrates that the mind is not analogous to a computer or other computational tools. However, a thorough analysis of both options will be briefly considered. In the case of (1), we are not able to understand our own functioning as we do not understand how we are able to access semantic content which is evident by our failures to transfer our semantic content to machines. In the case of (2), we are not able to understand our own functioning as we are machines with unsolvable problems (eg. why are we able to solve some problems but not others?). Thus, irrespective of siding with one of these perspectives over the other, both lead to the conclusion that we cannot understand our own functioning which begs the question of how we would be able to replicate this functioning (eg. imbed semantic understanding) in a machine.

In recognition and agreement with Gelgi's view that minds are essentially different and superior to machines, it is important to now explore why this difference exists, specifically, by explaining why the semantic content (eg. meaning) accessible to minds is not accessible to machines. In accordance with Penrose's view, human minds have access to semantic and syntactic content while machines only have access to syntactic content[5]. Further, this paper will argue that machine cognition functions by attempting to reduce semantic content into syntactic content. For example, artificial intelligence is being developed today to power self driving car algorithms. One task of this algorithm is to correctly identify street signs such as stop signs. However, a recent study found that if some tape is covering part of the stop sign, then the system incorrectly classifies it as a speed limit sign[4]. In contrast, a human would understand that it is a stop sign even with tape covering the sign. This is because humans have a semantic understanding of the stop sign whereas artificial intelligence only has a syntactic understanding of the stop sign's expected feature set, that is clearly easily fooled. This example illustrates the flaws in attempting to convert semantic content to syntactic content. Firstly, humans struggle to properly convert semantic content into digestible syntactic content even in basic cases such as the case of converting the meaning of a stop sign into a set of features necessary for artificial intelligence to identify it. This begs the question of whether and how we can convert more advanced semantic content into syntactic content such as truths that exist without proofs to a machine. Thus, "it is true, I believe that the meaning of an expression is a matter of how it is used. But this does not mean that meaning can be reduced to use. Rather, meaning permeates use. We manifest our understanding of expressions in our use of them"[3], and so this paper is in agreement with Moore's point that meaning (eg. semantic content) cannot be reduced to use (eg. syntactic content). Thirdly, the infinite nature of meaning conveyed by semantic content eludes itself from being reducible into the rigid containers of syntactic content. Moore notes "the meaning of an expression has infinite possibilities woven into it. Any expression can be applied in indefinitely many ways, for indefinitely many purposes, and to indefinitely many effects, whether literally, metaphorically, poetically, analogically, ironically, hyperbolically, precisely, roughly, or whatever"[3]. Thus, minds are essentially different and infinitely superior to machines because they are capable of processing semantic and syntactic content. In contrast, computers are only capable of processing syntactic content and the attempt to further enlighten machines by converting semantic content into it's language of syntactic content has largely failed.

In conclusion, this paper began by discussing the implications of Gödel's results in the field of mathematics, specifically, the realization that mathematics requires semantic understanding. Further, Gödel's results radiated to profoundly impacted our thoughts on thought itself, namely, human capacity for thought and machine capacity for thought. Gödel's results limit machine cognition by showing that general artificial intelligence will forever elude us. Additionally, an extension of his results show that our inability to understand our own mind is a further impediment to realizing an artificial intelligence that is equivalent to human intelligence. Finally, this paper has shown that minds are infinitely more powerful than machines because we have access to semantic and syntactic content while machines only have access to syntactic content. This paper has found that the attempt to bridge infinite semantic content into finite syntactic content necessary to achieve artificial intelligence equivalent to human intelligence constitutes an unsolvable problem. Thus, human intelligence remains infinitely superior to machine intelligence. In light of these conclusions, we can remain in wonder of the capacities of the human mind while seeking to appreciate, leverage and deploy the syntactical strengths of artificial intelligence.

References

1) Gelgi, F. (2004) Implications of Gödel's Incompleteness Theorem on A.I. vs. Mind, NeuroQuantology, Issue 3, 186-189, URL: http://www.neuroquantology.com/index.php/journal/article/download
2) Feferman (1995) *Kurt Gödel: Collected Works*, *III,*
3) Moore, A. W. (2019). *The infinite*. London: Routledge.
4) Schneier on Security. (n.d.). Retrieved from https://www.schneier.com/blog/archives/2017/08/confusing_self-.html
5) Smith, M. C. (n.d.). Godel and the incompleteness results.
6) Analysis. (2019, June 14). The Mind Can't Be Just a Computer. Retrieved from https://mindmatters.ai/2019/06/the-mind-cant-be-just-a-computer/