Gödel's Disjunctive Argument and Human Mind

Gabriel Chen[1*]

[1*]e-mail: gabrielchen@berkeley.edu.

Abstract. *In this paper, I discuss the concept of the first part of Gödel's disjunctive argument, and how Gödel think if the first part of his disjunctive argument is true, then the human mind cannot be reduced to the working of the brain.*
In the first part of this paper, I mainly address what the first part of Gödel's disjunctive argument is and how it connects to the human mind. And in the second part, I deliver my agreement on Gödel's opinion, and my belief on human mind cannot be converted into a Turing Machine.

In the lecture "Some basic theorems on the foundations of mathematics and their implications" Gödel gave in 1951, he addressed his famous disjunctive argument regarding mathematics, human mind, and Turing Machine as following.

> 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 unsolvabole diophantine[1] problems of the type specified.

Within the first part of this disjunctive, Gödel argued that all the evident axioms of mathematics cannot be written out into a finite series of sentences, or numbers, and this is the reason that mathematics is incomplete regardless any standpoint one takes. If this is true, the human mind surpasses any finite machines, more specifically, any Turing Machines. I will first discuss the argument Gödel made regarding how all evident axioms of mathematics cannot be written into a finite series of sentences, then derive the human mind surpasses any finite machines from there.

In order to write out all the evident axioms of mathematics, we first have to define what kind of sentences can be "axioms" which we would like to write out. This is

[1]The diophantine problem are of the following type: Let $P(x_1, \cdots, x_n, y_1, \cdots, y_m)$ be a polynomial with given integral coefficients and $n+m$ variables, $x_1, \cdots, x_n, y_1, \cdots, y_m$, and consider the variables x_i as the unknowns and the variables y_i as parameters, then the questions is, does the equation $P = 0$ have integral solutions for any integral values parameters, or are there any integral values of the parameters for which this equation has no integral solutions?

precisely defined as ”mathematics proper” by Gödel. These ”mathematics proper” are sentences which are absolutely true within not just a specific theory or system, but within any theory or system. For example, set theory holds its own truth no matter how. These sentences are unconditional truth within mathematics. And so is Logic, which is what we all agree prior defining anything else. On the other hand, there are conditional truth such as geometry or mechanics which is not always true within mathematics. Within a certain theory, unless we have some axioms to start with, or some mathematics proper to start with, we cannot have anything since we cannot derive anything from nothing.

It is very natural that these type of mathematics proper can be admitted differently from different standing point. There are two major different standing points while dealing with mathematics proper, one is the objective mathematics, and the other is the subjective mathematics. The objective mathematics defines mathematics proper as all true mathematical propositions within the system, and the subjective mathematics defines mathematics proper as all demonstrable mathematical propositions within the system. However, no matter which standpoint one believes a certain set of sentences are mathematics proper, mathematics is always inexhaustible. In a sense, the evident axioms can never be comprised in a finite rule regardless the standing point of mathematics proper as we will discuss next.

If we ever try to write down all evident axioms for mathematics, then what we precisely want to do is setting up a finite procedure, or a Turing Machine, which outputs all evident axioms of mathematics in a finite manner. However, we can never write out all axioms in a finite number of rules in the mean time having Gödel’s incompleteness theory. Given Gödel’s first incompleteness theory, ”whatever well-defined system of axioms and rules of inference may be chosen, there always exist diophantine problems of the type described which are undecidable by these axioms and rules, provided only that no false propositions of this type are derivable”, there is no such finite procedure which we can follow to write all axioms of a theory down since mathematics is such a well-defined system hence incomplete. On the other hand, given Gödel’s second incompleteness theory, ”for any well-defined system of axioms and rules, in particular, the proposition stating their consistency (or rather the equivalent number-theoretical proposition) is undemonstrable from these axioms and rules, provided these axioms and rules are consistent and suffice to derive a certain portion of the finitistic arithmetic of integers”, in other words, a well-defined system can never prove its own consistency within the system, and this is precisely at least one axiom we can not write it down, the consistency axiom. If one takes the standing point of objective mathematics, which is writing down all true mathematics propositions within the system, one can never write down the consistency axiom but the consistency is true. If one takes the standing point of subjective mathematics, which is writing down all demonstrable mathematical propositions, it is not guaranteed that this procedure is finite. Even there is a such process exists, one can never know that all propositions it outputs is correct because we cannot check it one by one once this process is infinite. Therefore, regardless the standing point one is holding for mathematics proper, the two Gödel's incompleteness theorems guarantee us that the evident axioms of mathematics can never be comprised in a finite rule.

For example, if we that we can finitely write done all mathematics proper, then we can axiomotize the set theory since all of the mathematics is reducible to abstract set theory as Gödel pointed out. Now our mission is reduced to axiomatizing set theory. However, the set of axioms of mathematics can be infinitely extended further and further, and in this way that all evident axioms of mathematics have no chance to be comprised in a finite rule producing them. Starting with the set of integers and define these integers as the "axioms of the first level", then apply the "set of" operation, which is taking the power set of our current set, over and over again on the set of integers to get next level of axioms. As long as the number of this kind of "set of" operation is not finite, this procedure can go on forever as many time as it wants. With this process, we can always get next level of axioms and we will never have an end point of iterating all of the axioms. Hence, mathematics can never be comprised into a finite rule in a sense of we can never finitely aximotize the set theory.

Then we proceed to how the evident axioms of mathematics cannot be comprised in a finite rule implies that the working of human mind can not be reduced to the working of human brain. But before we look into how this derive that human mind is not a finite machine as the work of human brain, we first have a look of how human brain works as referred by Gödel. Human brain, in a natural science way, is a physical object. It contains neurons and the connections between them. And how it works is purely based on chemical reaction and physics rules. Therefore, the pure working rule of human brain can be perfectly comprised into a finite rule. In other words, the working procedure of human brain can be modeled by a Turing Machine. However, this doesn't work the same for human mind. As mentioned above, all mathematics proper are sentences absolutely true within any system, which are sentenced admitted by human mind under any circumstances. In a sense, mathematics proper is fully understandable by human mind. Since we cannot comprise all mathematics proper into a finite rule, we cannot comprise human mind, which can understand all of them, in to a finite rule either. Hence, combining the argument that the working procedure of human brain can be modeled by a finite machine, and human mind cannot be comprised into a finite rule, the working of human mind cannot be reduced to the working of human brain.

Indeed, I do agree on the first part of Gödel's conjunctive argument that the working of human mind cannot be reduced to the working procedure of human brain, or any finite rules which can be proceeded by a Turing Machine. If the first part of Gödel's disjunctive argument is correct, this can be easily derived. However, this argument is facing a lot of challenges after decades of development on computability theory and neuron science.

In neuron science, the main challenging is understanding how human brain actually works. Given the first part of Gödel's disjunctive argument, we were assuming that the working process of human brain can be modeled using a Turing Machine since it only has finite number of neurons and the connections between them. As the working procedure of human brain follows certain natural science rules, no matter how large the set of rules is, it is finite, therefore all these rules can be modeled using a Turing Machine. The problem is that we have not fully understood the working procedure of human brain yet. If the working procedure of human brain does not work as we assumed, as a finite machine, we might get into trouble. On the other hand, regardless how the brain works, whether the working of human mind is the same as the working of human brain is still unknown.

In computability theory, although we prefer to believe the opposite, we still do not have a formal proof for $P = NP$ problem[2]. Which is whether all decision problems can be efficiently computed or not. If $P = NP$, all decision problems can be computed in polynomial time, including the working procedure of human mind, which is also a decision making process. On this problem, most people choose to believe that $P \neq NP$, which is there are some decision making problems cannot be efficiently computed, but we do not have any proof for this yet. If there is any one ever showed that $P = NP$, even though this is not very likely, we are getting into trouble of understanding the working process of human mind.

Obviously, these challenges of understanding the working of human mind are re-

[2] NP denote *nondeterministic polynomial*, it can be understood as the class of all search problem. P denote *polynomial*, it is the class of searching problems which can be solved in polynomial time. The $P = NP$ problem is asking whether all search problems can be solved efficiently in polynomial time ($P = NP$), or there exist some searching problems which can never be solved efficiently in polynomial time ($P \neq NP$).

ally hard to be answered directly. So, Instead of directly attacking these problems, which is not easy to achieve considering the limited resources we have right now, I will argue as following through a way that assuming human mind is a machine, then derive something very absurd and hard to believe which is true.

If the working of human mind can be indeed modeled using a Turing Machine, it means every thought a human can ever give out can be modeled within a finitely axiomotized system. More specifically, we can build a thinking machine, and with some finite number of axioms encoded inside of it, it can derive every thought a human mind can ever give out. In other words, there exist a machine which can simulize human thoughts. if this were true, we can build a machine to derive all mathematics proofs since all proofs are basically human thoughts. Then there is no need for mathematicians or human intellegent. It looks absurd already, but if we look more closely, we can derive $P = NP$ from here. Since we can enumerate all the possibilities of human thoughts if human mind is a machine, then we can first enumerate everything and loop through all the possibilities for a searching problem, which apparently, having a polynomial running time. Then, things we can derive from $P = NP$ would raise a lot of problems. For example, it is as easy to write any article as to write a good article, it is as easy to act as to become an actor, and it is as easy to learn from a math book as to come up with all proof by oneself. This is too absurd to be true.

Overall, given the first part of Gödel's disjunctive argument, mathematics proper cannot be finitely written out as a finite number of axioms. From here, since human mind can fully understand all mathematics proper regardless the standing point, either objective mathematics or subjective mathematics, it derives that the working process of human mind cannot be reduced to the working process of human brain, which can be modeled as a Turing Machine with finite number of procedures as referred by Gödel. In a sense, human mind infinitely surpass the power of any finite machines. On the other hand, if the working procedure of human mind can be modeled as finite machine, the result can be derived would be too absurd to believe.