PHILOSOPHICAL CONSEQUENCES OF GÖDEL’S THEOREM

I. General consequences of philosophical importance

Gödel’s theorem can be stated as follows: for every formalized axiomatic theory $S$ which contains a sufficient portion $Ar$ of elementary arithmetic there exists a sentence $G_S$ such that

(*) if $S$ is consistent then $G_S$ is true but unprovable in $S$ (the 1st Gödel’s theorem) and, moreover, this property is enjoyed by the sentence $\text{Con}_S$, a natural formalization of the statement of consistency of the theory $S$ (the 2nd Gödel’s theorem).

1. It seems certain that finite methods introduced by the formalist school, based on immediate intuition of symbols, are expressible in $Ar$. If it is so then the 2nd theorem shows that these methods are not sufficient for proving the consistency of $S$. Thus the original Hilbert program fails.

2. Limitations of the axiomatic method: Gödel’s theorem shows that no formal system embodies the whole of mathematics, even of elementary number theory, and even the domain of diophantine equations. The least statement was shown in 1970: $G_S$ may be taken of the shape “there is no solution for $E_S$”, where $E_S$ is a suitable diophantine equation. In other words: no finite description of our universe (conceived as including the ideal, or mental, entities like the natural numbers) is possible.
3. No machine can produce all, and only, truths. Even the truths about the existence of solutions of diophantine equations. The reason: it is known that the output of each machine may be considered as a formalized.

4. In Gödel’s own words: “Human mind cannot formalize all its mathematical intuitions”. Because if we could then a new intuition would appear transcending the formalization, namely the formalization’s consistency (we assume here it is consistent).

5. Does it follows from Gödel’s theorem (leaving other arguments aside) that truth is different from (non-equivalent to, non-reducible to) provability? The answer is ‘yes’ if we mean provability in a single system. Still, truth may be provability in an informal sense, as maintain intuitionists, or truth may be given by provability in an informal collection of formal systems (Wang).

6. It seems that all our explanations how to use the term ‘natural number’ can be described by a formal system. So some truths remain outside, and we can see they are truth since the meaning of the term is clear enough. Is this meaning (concept) innate? The example shows that meaning cannot be explained by use.

The least examples cannot be decided in a conclusive manner, mathematical theorems are but one face of philosophical problems. There is however an exception: it has been argued that Gödel’s theorem implies that “human minds are not machines”. The usual attempts fail as we show below. Gödel himself was of the same opinion, even though he was looking for the arguments that “the laws of thought are not mechanical” too.

II. Is there a mathematical proof that we are not machines?

A well known exposition of the positive answer is due to Lucas (Philosophy 36, 1961). If Mechanist proposes a machine $M$ saying that it is equivalent to the mathematical powers of Lucas then Lucas produces the Gödelian formula $G_M$ corresponding to the theory $T(M)$ consisting of arithmetical statements provable by $M$. Now either (Case 1) $T(M)$ is consistent or (Case 2) $T(M)$ is inconsistent. In the first case $G_M$ is unprovable by $M$.
but seen to be true, i.e. provable, by Lucas. In the second case everything can be proved by $M$ (with the help of the classical logic), which is not true about Lucas who certainly won’t prove $2 + 2 = 5$. In any case Lucas can “out-Gödel” $M$, i.e. show his being different from $M$.

In principle the above procedure can be mechanized. Let us assume that all machines are listed in the sequences $M_1, M_2, \ldots$ (the index corresponds, usually, to a description of the machine). There exists a recursive function $g$ such that for every $n$

$$ (+ \text{ if } T(M_n) \text{ is consistent then } g(n) \text{ is (the Gödel number of) a (Gödelian)} \text{ formula which is true and unprovable in } T(M_n).$$

Objections (not necessarily made by those who believe that man is machine, or his mathematical powers are):

1. We can imagine a machine simulating the “out-Gödeling” (in fact (+) is provable in fairly weak theories). Machines are no worse than we are. The reply is that this machine can be out-Gödeled too. In general, the aim is not to dominate all machines at one bur rather each machine proposed by Mechanist. Lucas describes the procedure as dialectical, or as a game in which Lucas wins in every move.

2. It has been argued that the procedure does not work if the machine’s program is not known. Lucas, at least as far as mathematical powers are considered, may be equal to a machine without being able to find $n$ in order to produce $g(n)$. The reply is that theoretically it is possible to do this. The appropriate Gödelian sentence is somewhere there, waiting for us. And we know it is true provided the machine under consideration is consistent (i.e. $T(M)$ is consistent). But how do we know that a machine is consistent? Even if it is we may be unable to prove this. Rogers, Wang, Putman and others stressed the fact that all we know is (+) or (∗), a conditional statement. It provides the argument in Case 1, but how do we know, whether we are in Case 1 or Case 2? If Mechanist presents a machine Lucas may be unable to decide its consistency. In this case it is not so simple to say that it is theoretically possible. The question “is a given machine consistent?” is recursively undecidable, the set $C = \{ n : T(M_n) \text{ is consistent} \}$ is not recursive. Still, theoretically either Case 1 or Case 2 applies, so the whole argument may be seen as valid, only its dialectical character cannot be retained.

3. We have to consider the problem: “Are we really consistent?”.
Case 2 of the argument depends on the positive answer. But the 2nd Gödel’s theorem, seems to imply that if we are consistent we cannot prove this mathematically! The reason is that we prove mathematically can be simulated by a machine expressing a part of our mathematical powers, so the machine would prove its own consistency.

Everyday experience shows that humans are rather inconsistent. Lucas remarks: “Certainly women are, and politicians”. Putman maintained that it is possible that we are inconsistent machines. Apparently inconsistent machines prove everything. However we can only say that everything follows, not that an inconsistent machine actually produces all formulas. This is a weak point of Lucas’ proof.

Lucas’ reply is that our inconsistency is rather like machine malfunctions. We remove reasons causing inconsistency. And to “fallible but self correcting” machine the “out-Gödeling” procedure may be applied, that is, its Case 1.

I agree that we have to believe that mathematics is consistent. Consistency is like a regulative idea for our mathematical activities. So we assume our fundamental consistency. But it’s not the end of the analysis.

4. The final blow to the Lucas’ argument was given by G. Lee Bowie (jour. Philosophical Logic 11, 1982). He remarked that, whatever Lucas may claim about his consistency, we know enough to prove that he is inconsistent. As we know the systematic use of the Lucas’ procedure can be expressed as the recursivity of the appropriate function \( g \). The point is that the range of \( g \), i.e. the set \( A = \{ g(n) : n = 1, 2, \ldots \} \), is inconsistent! This is proved as follows: The set \( A \) is equal \( T(M_k) \) for a suitable \( k \). If it were consistent, then by virtue of (+) \( g(k) \) is unprovable in \( T(M_k) \). But \( g(k) \) belongs to \( A \), that is to \( T(M_k) \), which is a contradiction.

Considering the whole of the range of \( g \) is adequate, in particular in the light of the remark at the end of II.2. One might attempt to save Lucas by applying \( g \) only to consistent machines. This is false: there is no decision procedure for testing the consistency (the set \( C \) is not recursive). Moreover, the set \( C \) cannot be effectively generated (it is not r.e.). Whatever modifications Lucas might do, as long as the modified procedure is a single procedure described fully in a finite way, that it is an effective procedure, the above proof is valid. Let us observe that the Lee Bowie’s proof requires only that 1° \( g \) is partial recursive, 2° the domain of \( g \) includes the set \( C \) (of consistent machines), 3° for \( n \) in \( C \) the formula (with the Gödel number)
$g(n)$ is unprovable in $T(M_n)$. And the conditions $1^o, 2^o, 3^o$ are all essential for the idea of “out-Gödeling” the Mechanist (observe that we don’t need the condition that $g(n)$ is true).

It has been proved mathematically that the class of inconsistent humans is certainly not empty: it contains the analytic philosophers believing in Gödel based mathematical proof of their own superiority over machines.

Of course, we feel we are consistent. This may be misleading, as the story of Lucas shows, but the feeling cannot be lightly dismissed. Assuming our consistency makes it possible to hold a view Gödel himself held (according to Wang). Namely, it is not excluded man intuition (at least in the domain of mathematics) and the $M$ is even empirically discoverable. We could not, however, prove this equivalence. How can such a machine emerge? We may apply the concept of evolution to machines, which was done after von Neumann had observed that, in principle, self-reproducing machines can exist. Rucker described in a colorful way a civilization of computers on the moon. It can be programmed with the use of random devices that new generations of computers differ slightly from the previous ones. Having introduced mutations we can imagine that natural selection and evolution of the machines takes place. In this way a computer may come to existence that is equivalent to Lucas but no man would be able to detect this.

Finally, I must avow that I believe that we are no machines. The best argument I know (except the theological ones that are not universally acceptable) is based on our selfconsciousness. This has, however, little to do with mathematical intuitions. Gödel’s theorem provides another argument under the assumption that we are able to solve every mathematical problem. In fact, to draw the conclusion it is sufficient to believe that we can solve every diophantine equation. We know that no machine can do this. The belief in our ability to solve every mathematical problem was stated by Hilbert. It was shared by Gödel.