WHAT GÖDEL’S THEOREM REALLY PROVES

Antonio Leon
Interciencia, Salamanca, Spain
interciencia.es

Abstract. It is proved in this paper the undecidable formula involved in Gödel’s first incompleteness theorem would be inconsistent if the formal system where it is defined were complete. So, before proving the formula is undecidable it is necessary to assume the system is not complete in order to ensure the formula is not inconsistent. Consequently, Gödel proof does not prove the formal system is incomplete but that, once assumed it is incomplete, it is possible to define an undecidable formula within the system. This conclusion makes Gödel’s incompleteness theorems devoid of substance.

Contents

INTRODUCTION: NATURE, LANGUAGE AND LOGIC 1
PRELIMINARY CONVENTIONS AND DEFINITIONS 2
THE FIRST LAW OF LOGIC 4
THEOREM OF THE INCONSISTENT SUBJECT 6
CONSEQUENCES ON SOME SEMANTIC PARADOXES 7
CONSEQUENCES ON GÖDEL’S THEOREM 9
SHORT EPILOG 13
References 13

Introduction: nature, language and logic

1. Gödel’s first incompleteness theorem (Gödel’s theorem hereafter) was inspired by the Liar Paradox. In fact, as Gödel himself recognized [11, p. 56], it was this paradox what leads him, via Richard paradox, to his celebrated theorem. In the place of a sentence that asserts of itself that it is not true, now it is an abstract formula \( G \) of which we know, through its ordinary language interpretation, it asserts of itself that it is not demonstrable in the formal system \( P \) where it is defined. Therefore, \( G \) is an abstract meta-mathematical statement whose veracity can be tested via ordinary logic (the basic logic subsumed into ordinary thinking). It is in this way that ordinary logic and the logic of an abstract formal system meet in Gödel’s theorem.

2. Here we will analyze the encounter from the perspective of statement logic which is directly founded on ordinary logic. Although before doing it, some considerations on the relations between the formal consistency of the physical world and the origin of ordinary logic will be made. The objective of these considerations is to perform a brief evaluation of the relevance of the statement logic scenario where the ordinary language interpretation of Gödel’s theorem will be reanalyzed.
3. Nature seems to be consistent, i.e. subjected to invariable rules. Although most of those rules have statistical grounds, at the mesoscopic level where life emerges, nature appears to be clearly consistent. Living beings evolve and thrive in syntony with that formal consistency. This means we have been forced, in evolutionary terms, to grasp the logic of nature: we need to know the way nature works to perform the appropriate responses in order to survive and reproduce [15, pp. 11-24]. In addition, we, human beings, are totally confident in the formal consistency of the physical world in order to develop our science and technology, and then to know and predict the behaviour of nature as well as to prospect and exploit all type of material and energetic resources and to invent and built all types of devices, machines, edifices etc. The success in achieving all those objectives suggests that nature is, in effect, formally consistent, subjected to invariable rules. No exception is known to this general conclusion.

4. In accordance with 3, our natural way of reasoning (ordinary logic) is a consequence of the consistent behaviour of nature. It has been modeled by organic evolution in consonance with the rationality of the physical world. The same world we try to describe, explain and exploit. As just noted, the productivity of our modern science and technology makes it evident that, in fact, our ordinary logic is in syntony with the logic of the physical world. In this sense, let us recall that formal sciences, experimental sciences, human sciences and ordinary logic share, all of them, at least two universal laws: the First and the Second Laws of logic. In some way, therefore, those laws must reflect, at the most basic level, the formal consistency of nature.

5. It seems reasonable to conclude, therefore, that ordinary logic, the logic of ordinary thinking, is not arbitrary but grounded on the logic of the physical world. Ordinary logic, in turn, inspired the birth of formal logic, the birth of logic as a scientific discipline. Formal logic improves ordinary logic by concentrating on form up to the point of becoming independent of the content that characterizes ordinary language. Notwithstanding, in most of the cases, abstract formulas of formal systems can be interpreted in terms of ordinary language, what makes it possible to check their veracity and consistency from the perspective of statement logic. As we will see, this is the case of the above mentioned Gödel’s formula $G$.

**Preliminary conventions and definitions**

6. The discussion that follows makes use of the usual logic symbols $\neg$ (negation); $\Rightarrow$, $\downarrow$ (implication); $=$ (equality). When needed, formulas and statements will be

\footnote{We now suspect the laws of physics we are referring to, are laws of collective behaviour emerging from other more basic level of reality governed by other more basic set of rules [13], [14]}
named by letters \((p, q, r, \ldots)\). Monadic self-referring statements \(p\) will be written in the form:

\[
p: \ p \text{ is } X
\]

where \(X\) is any predicate of \(p\). As usual, equations will be referred to by numbers in brackets. Paragraphs will be referred to by the numbers, without brackets, that appear at the beginning of each one of them.

7. A formal system will be understood as:
   - A set of symbols.
   - A set of rules that state the way of forming strings of symbols (well formed formulas).
   - A set of axioms (well formed formulas).
   - A set of inference rules.

A formal system will be said consistent if for every well formed formula \(F\) proved within the system, the formula \(\neg F\) cannot be proved within the system (or alternatively if there is a formula in the system that cannot be proved within the system). Here, we will only deal with consistent formal systems (formal systems for short) and well formed formulas. Finally, a formal system will be said complete if for every well formed formula \(F\), either \(F\) or \(\neg F\) can be proved within the system. If the system is not complete it will be said incomplete.

8. We also assume here that, with respect to the formal system \(P\) where it is defined, a formula \(F\) can be:
   - Demonstrable: \(F\) is derived from \(P\)’s axioms by application of \(P\)’s inference rules.
   - Non-demonstrable: \(\neg F\) is demonstrable in \(P\).
   - Undecidable: Neither \(F\) nor \(\neg F\) are demonstrable in \(P\).
   - Inconsistent: \(F\) implies a \(P\)’s formula and its negation.

9. As usual in statement logic, we will only consider well formed declarative sentences, i.e. sentences capable of being true or false. They will also be referred to as statements. In addition, all statements will be monadic, with only one subject and only one predicate, and being the predicate anything that can be said on the subject, even a predicate that predicates itself. We will assume a monadic statement can be:
   - True: The case is what the statement asserts to be.
   - Non true (false): The case is not what the statement asserts to be.
   - Undecidable: True or False but impossible to decide which is the case.
   - Inconsistent: It implies a contradiction, i.e a statement and its negation.

10. Evidently true statements and formulas cannot be inconsistent. Otherwise a truth would imply a falsehood through one of the terms of the involved contradiction.
Furthermore, Gödel’s theorem would not be necessary because we would have true formulas that could not be proved within consistent systems.

11. If $p$ is a well formed declarative sentence and $X$ a predicate of $p$ then it will be assumed that:

11-1) The following three sentences:

$$\neg p; \quad p \text{ is } X; \quad p \text{ is not } X$$

are all of them well formed declarative sentences.

11-2) $p$ is not $X = \neg(p \text{ is } X)$

11-3) $\neg(\neg p) = p$

being all of them standard rules in all formal systems, including statement logic.

12. The subject of a monadic sentence may be another sentence, a predicate, a noun, a formula or anything else capable of being predicated. The subject $S$ of a monadic sentence of the type:

$$S \text{ is } X$$

will be said to be a subject of consistency of the predicate $X$ if it takes one, and only one, of the following four values:

12-1) $X$

12-2) $\neg X$

12-3) Undecidable: $X$ or $\neg X$ but impossible to decide which is the case.

12-4) Inconsistent: Both $[S \text{ is } X]$ and $[S \text{ is not } X]$ lead to contradictions.

The First Law of Logic

13. As is well known at least since Aristotle’s time [1], [2], the formal structure of all sciences rests on two fundamental assumptions: the laws of logic. The next discussions are mainly concerned with the first of those laws.

14. The First Law of logic is usually stated in informal terms as: 'something is what it is, and not what it is not' (Aristotle); or as $A = A$, and the like. From the statement logic point of view, the First Law can be written as ([3, p. 139]):

$$p \Rightarrow p$$

where $p$ is any statement. Implication (3) translates the sense of identity to the world of statements. Notice that, in fact, implication (3) is always true, it is a tautology, a law.

---

2 Or three if we include the Law of the Excluded Middle.
15. As we will see in the next section, the First Law plays a capital role in the proof of the Theorem of the Inconsistent Subject. This theorem applies to subjects of consistency $S$ of a predicate $X$ for which it holds:

$$S \text{ is } X \Rightarrow S \text{ is not } X$$

(4)

$$S \text{ is not } X \Rightarrow S \text{ is } X$$

(5)

16. A well known example of statement satisfying (4) and (5) is the Liar statement $p$ [4] [18], in which $X$ is the predicate $\text{true}$:

$$p: \text{p is not true}$$

(6)

It is quite clear that if $p$ is true, then it is true what $p$ asserts; and being what $p$ asserts that $p$ is not true, then $p$ is not true:

$$p \text{ is true } \Rightarrow \text{p is not true}$$

(7)

Alternatively, if $p$ is not true then it is not true what $p$ asserts; and being what $p$ asserts that $p$ is not true, then it is not true that $p$ is not true, i.e $p$ is true:

$$p \text{ is not true } \Rightarrow \text{p is true}$$

(8)

Figure 1. An informal version of the Liar paradox. Could this man make it explicit of what he is predicating its falsity? If I were that man I must confess I could not.

17. Other paradoxes, as Richard paradox [17], Grelling-Nelson paradox [6], [12], Russell’s paradoxes [16] [9] [5] [7] share with the Liar paradox some suspicious features:

17-1) They have self-referring subjects of consistency.

17-2) They have a (direct or indirect) negative form: to be non-$X$. Where $X$ may be: true; richardian; heterologic; to predicate itself; to shave himself; to belong to itself.

17-3) $X$ is sensitive to the double negation: non-(non-$X$) = $X$

17-4) The involved statements are not empirically verifiable, so we must speculate on whether their corresponding subjects are, or not, $X$.

\[\text{The liar paradox can be expressed in many different ways, even in the form of several circularly related statements.}\]
Theorem of the Inconsistent Subject

18. As we will see, the next theorem is an efficient instrument to unmask the inconsistent nature of the above mentioned paradoxes. It is also the instrument we will make use of to reanalyze the ordinary language interpretation of Gödel undecidable formula.

Theorem of the Inconsistent Subject.-Let $S$ be any subject of consistency of the predicate $X$. If $S$ and $X$ satisfy:

$$ S \text{ is } X \Rightarrow S \text{ is not } X \quad (9) $$

$$ S \text{ is not } X \Rightarrow S \text{ is } X \quad (10) $$

then $S$ is an inconsistent subject of the predicate $X$

19. Proof.-Consider the statements $p$ and $q$:

$$ p: \ S \text{ is } X \quad (11) $$

$$ q: \ S \text{ is not } X \quad (12) $$

Evidently:

$$ p = \neg q \quad (13) $$

$$ q = \neg p \quad (14) $$

Now (9) and (10) can be rewritten as:

$$ p \Rightarrow q \quad (15) $$

$$ q \Rightarrow p \quad (16) $$

With respect to $p$, and according to the First Law of logic, we have:

$$ p \Rightarrow p \quad (17) $$

which in accord with (13), can also be written as:

$$ p \Rightarrow \neg q \quad (18) $$

Thus, by (15) and (18), we have:

$$ p \Rightarrow q \quad \left\{ \begin{array}{l}
          p \Rightarrow \neg q \\
        \end{array} \right. \quad (19) $$

Therefore, $p$ is inconsistent. Similarly, with respect to $q$ and according again to the First Law, we have:

$$ q \Rightarrow q \quad (20) $$

which in accord with (14), can also be written as:

$$ q \Rightarrow \neg p \quad (21) $$

Thus, by (16) and (21), we have:

$$ q \Rightarrow p \quad \left\{ \begin{array}{l}
          q \Rightarrow \neg p \\
        \end{array} \right. \quad (22) $$

Therefore, $q$ is inconsistent.
20. Once proved $p$ and $q$ are inconsistent let us examine the possibilities for $S$:

20-1) If $S$ were $X$ then $p$ would be true, what is impossible because $p$ is inconsistent.

20-2) If $S$ were $\neg X$ then $q$ would be true, what is impossible because $q$ is inconsistent.

20-3) If $S$ were undecidable either $p$ or $q$ would be true (though we could not decide which is the case), what is impossible because we know both of them are inconsistent.

In consequence, and taking into account that, with respect to the predicate $X$, the subject of consistency $S$ must take one, and only one, of the four values: $X$, $\neg X$, undecidable or inconsistent, the subject $S$ can only be an inconsistent subject of the predicate $X$. This is also the conclusion that directly follows from the proved facts that $p$ [$S$ is $X$] and $q$ [$S$ is not $X$] lead to the contradictions (19) and (22) respectively. $S$ is, therefore, an inconsistent subject of the predicate $X$, as the Theorem of the Inconsistent Subject states.

21. It is worth noting the above proof of the Theorem of the Inconsistent Subject is exclusively based on the First and the Second law of logic. Also noteworthy is the fact that all substitutions and inferences rules used in the proof are absolutely elementary and universal in all logic systems, from abstract formal systems, as Gödel’s formal system $P$ (see below), to statement logic and ordinary logic. It is, therefore, a basic theorem of logic.

22. An immediate consequence of the above theorem is the following:

**Corollary.**-If $S$ is an inconsistent subject of the predicate $X$, then the statements:

\begin{align*}
S & \text{ is } X \quad (23) \\
S & \text{ is not } X \quad (24)
\end{align*}

are both inconsistent statements.

**Proof.**-Contradictions (19) and (22) prove respectively that statements (23) and (24) are inconsistent.

**Consequences on some semantic paradoxes**

23. Before examining the consequences of the Theorem of the Inconsistent Subject on Gödel formula, and with the objective of illustrating its broad scope, we will examine the consequences of this theorem on some well known semantic paradoxes. To begin with, consider the Liar paradox:

\[ p: \text{p is not true} \quad (25) \]

As we saw in 16, it holds:
Therefore, and in accord with the Theorem of the Inconsistent Subject, the statement $p$ is an inconsistent subject of the predicate $true$, and the statements $[p \text{ is true}]$ and $[p \text{ is not true}]$ are inconsistent.

24. Let us consider Grelling-Nelson paradox:

$$\text{Heterologic is not heterologic}$$  \hspace{1cm} (27)

where heterologic ($H$ for short) is an adjective that does not describe itself, being non-heterologic (autologic) if it does. On the one hand we have:

$$H \text{ is not } H \Rightarrow H \text{ is } H$$

because if heterologic is not heterologic then it describes itself: heterologic is heterologic. On the other hand we also have:

$$H \text{ is } H \Rightarrow H \text{ is not } H$$

because if heterologic is heterologic then it does not describe itself, i.e. it is not heterologic. Thus, in accord with the Theorem of the Inconsistent Subject, heterologic is an inconsistent subject of the predicate heterologic, and the statements $[H \text{ is not } H]$ and $[H \text{ is } H]$ are both inconsistent.

25. Consider now Richard paradox:

$$k \text{ is not richardian}$$  \hspace{1cm} (28)

where $k$ is the index of the property of being richardian$^4$ ($R$ for short) in the indexed list of properties of natural numbers. On the one hand we have:

$$k \text{ is not } R \Rightarrow k \text{ is } R$$  \hspace{1cm} (29)

because if $k$ is not richardian then $k$ satisfies the property it indexes, which is the property of being richardian. On the other hand we also have:

$$k \text{ is } R \Rightarrow k \text{ is not } R$$  \hspace{1cm} (30)

because if $k$ is richardian then $k$ does not satisfy the property it indexes, which is the property of being richardian. Thus, in accord with the Theorem of the Inconsistent Subject, $k$ is an inconsistent subject of the predicate richardian, and the statements $[k \text{ is not } R]$ and $[k \text{ is } R]$ are both inconsistent.

26. A more general case is the paradox of Russell’s predicate [5]:

$$\text{Russellian is not russellian}$$  \hspace{1cm} (31)

$^4$An index is richardian if it does not satisfy the property it indexes, and non-richardian if it does.
where russellian (R for short) is a predicate that does not predicate itself, being non-russellian if it does. On the one hand we have:

\[ R \text{ is not } R \Rightarrow R \text{ is } R \]

because if russellian is not russellian then it predicates itself: russellian is russellian. On the other hand we also have:

\[ R \text{ is } R \Rightarrow R \text{ is not } R \]

because if russellian is russellian then it does not predicate itself, and consequently it is not russellian. Thus, in accord with the Theorem of the Inconsistent Subject, russellian is an inconsistent subject of the predicate russellian, and the statements \([R \text{ is not } R]\) and \([R \text{ is } R]\) are both inconsistent.

27. In a similar way, it can also be proved that:

27-1) If \(S\) is Russell’s superbarber that shaves all those, and only those, that do not shave themselves, then \(S\) is an inconsistent subject of the predicate to shave himself, and the statements \([S \text{ shaves himself}]\) and \([S \text{ does not shave himself}]\) are inconsistent.

27-2) If \(R\) is the set of all sets that do not belong to themselves then \(R\) is an inconsistent subject of the predicate to belong to itself, and the statements \([R \text{ belongs to itself}]\) and \([R \text{ does not belong to itself}]\) are inconsistent.

Figure 2. Left.-To shave or not to shave himself, that is the question for Russell’s superbarber. Right.-Russell’s set trying to belonging to itself.

Consequences on Gödel’s theorem

28. As is well known, Gödel’s theorem solved the metamathematical deficiency of Richard’s paradox: the inclusion of metamathematical statements in the formal system. Gödel formula is, in fact, a meta-mathematical statement interpretable in terms of ordinary language. As Richard statement, this one is also self-referring, has a negative form and is not empirically verifiable.

\(^5\)Published in 1931 [10], [11]
29. In his paper of 1931, Gödel proved the existence of a formula $G$ in a formal system $P$ such that if $G$ is $P$-demonstrable (P-dem for short) then $\neg G$ is also $P$-dem. And if $\neg G$ is $P$-dem then $G$ is also $P$-dem. Therefore, if $P$ is consistent then $G$ is undecidable and, in consequence, $P$ is incomplete. This is the essential content of his first incompleteness theorem (Theorem VI of his paper). Once interpreted in terms of ordinary language, $G$ is a self-referring metamathematical statement that asserts of itself it is not $P$-dem:

$$G: \quad \text{G is not P-dem} \quad (32)$$

And being undecidable, $G$ is not, in fact, $P$-dem. So $G$ is true.

30. Assume, only for a moment, that $P$ were a complete formal system, i.e. a formal system such that for every well formed formula $F$ of $P$, either $F$ is $P$-dem or $\neg F$ is $P$-dem. Under this hypothesis we will prove that:

30-1) $G$ is an inconsistent subject of the predicate $P$-dem.

30-2) $G$ is an inconsistent subject of the predicate true

31. Consider the ordinary language interpretation (32) of Gödel’s formula $G$. Under the assumption that $P$ is complete, we can write:

$$G \text{ is not } P\text{-dem}$$

$\Downarrow$ (By completeness)

$\neg G \text{ is } P\text{-dem}$

$\Downarrow$ (Gödel proof)

$$G \text{ is } P\text{-dem}$$

And:

$$G \text{ is } P\text{-dem}$$

$\Downarrow$ (Gödel proof)

$\neg G \text{ is } P\text{-dem}$

$\Downarrow$ (By consistency)

$$G \text{ is not } P\text{-dem}$$

Then, we have:

---

$^6$An adaptation of Russell-Whitehead’s Principia Mathematica by means of which it is possible to establish the theorems of whole-number arithmetic.
In consequence, and according to the Theorem of the Inconsistent Subject, $G$ is an inconsistent subject of the predicate $P$-dem.

32. Under the same hypothesis that $P$ were a complete formal system, we can also write:

\[
\begin{align*}
G \text{ is not true} \quad &\Downarrow \\
\Rightarrow G \text{ is P-dem} \quad &\Downarrow \quad \text{(By Gödel proof)} \\
\neg G \text{ is P-dem} \quad &\Downarrow \quad \text{(By consistency)} \\
G \text{ is not P-dem} \quad &\Downarrow
\end{align*}
\]

On the other hand:

\[
\begin{align*}
G \text{ is true} \quad &\Downarrow \\
\Rightarrow G \text{ is not P-dem} \quad &\Downarrow \quad \text{(By completeness)} \\
\neg G \text{ is P-dem} \quad &\Downarrow \quad \text{(By Gödel proof)} \\
G \text{ is P-dem} \quad &\Downarrow
\end{align*}
\]

Then, we have:

\[
\begin{align*}
G \text{ is not true} \Rightarrow G \text{ is true} \cr
G \text{ is true} \Rightarrow G \text{ is not true}
\end{align*}
\]

In consequence, and according to the Theorem of the Inconsistent Subject, $G$ is an inconsistent subject of the predicate true.

33. We have just proved that under the hypothesis that $P$ were complete, the ordinary language interpretation of $G$:

\[
G: G \text{ is not P-dem}
\]
satisfies the conditions (9)-(10) of the Theorem of the Inconsistent Subject for the predicates $P\text{-}dem$ and $true$. Consequently, and in agreement with Corollary 22, the self-referring sentences:

\[
\begin{align*}
G &: G \text{ is true} \\
G &: G \text{ is not true} \\
G &: G \text{ is } P\text{-}dem \\
G &: G \text{ is not } P\text{-}dem
\end{align*}
\]

are all of them inconsistent. Therefore, the ordinary language interpretation (35) of Gödel formula $G$ is inconsistent. Now then, having been derived with the only aid of the First and the Second Laws of logic (proof of the Theorem of the Inconsistent Subject), the inconsistency of the ordinary language interpretation of $G$ must also apply to the abstract formula $G$ in the formal system $P$ because:

33-A) The First and the Second Laws of logic are axioms of Gödel’s system $P$.
33-B) The First and the Second Laws of logic are the same for all logical systems.
33-C) Gödel’s abstract formula $G$ can be interpreted in terms of ordinary language.
33-D) The ordinary language interpretation (35) of Gödel’s formula is the correct one.
33-E) The ordinary language interpretation (35) of Gödel’s formula is inconsistent.

Therefore, it holds the following:

**Theorem of completeness 1.**-If Gödel’s formal system $P$ were complete then the undecidable formula $G$ would be inconsistent.

34. Theorem of completeness 1 means that if we assume the hypothesis that all formal systems are complete, then Gödel abstract formula would not serve to disprove that hypothesis, simply because that formula would be inconsistent in any complete formal system. Or in other words, in any *supposedly* complete formal system, Gödel abstract formula would not serve as a counterexample to prove that system is not complete because that formula would be inconsistent in that system. What Gödel’s incompleteness theorem really proves is that a formula that is inconsistent in a complete formal system would be undecidable in a *supposedly* incomplete formal system.

35. Let us call Gödel’s incompleteness condition to the following one:

A sufficient condition for a formal system $P$ to be incomplete is the existence of a true formula $F$ in $P$ such that if $F$ is $P\text{-}dem$ then $\neg F$ is also $P\text{-}dem$; and if $\neg F$ is $P\text{-}dem$ then $F$ is also $P\text{-}dem$. 

![Figure 3. Informal language interpretation of Gödel formula.](image)
As a consequence of the Theorem of the Inconsistent Subject and from the above proofs 31 and 32, it holds the following:

**Theorem of completeness 2.** Gödel’s incompleteness condition does not prove a formal system is incomplete.

**SHORT EPILOG**

36. I agree with Galileo’s opinion on the Liar Paradox [8, pp 93-94]:

[...] in this type of sophisms, you would be hanging around forever without ever reaching a conclusion.

In my opinion all self-referring sentences and formulas are childish and capricious games of words and symbols that have made us lose a lot of time and money. Until now they have been absolutely useless in order to explain the physical world: self-reference has no reflection in that physical world. We, human beings, are the only known natural objects with the ability to refer to other objects, including ourselves. Recall, on the other hand, that self-reference had to be axiomatically removed from set theory to avoid certain persistent inconsistencies. I think that, as Wittgenstein suggested [19, 3.332, p. 43], ordinary language and formal languages should follow the example of set theory.

References

What Gödel’s theorem really proves