Liar paradox

From Wikipedia, the free encyclopedia

In philosophy and logic, the classical liar paradox or liar's paradox is the statement of a liar who states that they are lying: for instance, declaring that "I am lying" or "everything I say is false". If they are indeed lying, they are telling the truth, which means they are lying. In "this sentence is a lie" the paradox is strengthened in order to make it amenable to more rigorous logical analysis. It is still generally called the "liar paradox" although abstraction is made precisely from the liar himself. Trying to assign to this statement, the strengthened liar, a classical binary truth value leads to a contradiction.

If "this sentence is false" is true, then the sentence is false, but if the sentence states that it is false, and it is false, then it must be true, and so on.

History[edit]

The Epimenides paradox (circa 600 BC) has been suggested as an example of the liar paradox, but they are not logically equivalent. The semi-mythical seer Epimenides, a Cretan, reportedly stated that "All Cretans are liars."[1] However, Epimenides' statement that all Cretans are liars can be resolved as false, given that he knows of at least one other Cretan who does not lie. It is precisely in order to avoid uncertainties deriving from the human factor and from fuzzy concepts that modern logicians proposed a "strengthened" liar such as the sentence "this sentence is false".

The paradox's name translates as pseudómenos lógos (ψευδόμενος λόγος) in Ancient Greek. One version of the liar paradox is attributed to the Greek philosopher Eubulides of Miletus who lived in the 4th century BC. Eubulides reportedly asked, "A man says that he is lying. Is what he says true or false?"[2]

The paradox was once discussed by St. Jerome in a sermon:

"'I said in my alarm, Every man is a liar!' Is David telling the truth or is he lying? If it is true that every man is a liar, and David's statement, "Every man is a liar" is true, then David also is lying; he, too, is a man. But if he, too, is lying, his statement: "Every man is a liar," consequently is not true. Whatever way you turn the proposition, the conclusion is a contradiction. Since David himself is a man, it follows that he also is lying; but if he is lying because every man is a liar, his lying is of a different sort."[3]

The Indian grammarian-philosopher Bhartrhari (late fifth century AD) was well aware of a liar paradox which he formulated as "everything I am saying is false" (sarvam mithyā bravīmi). He analyzes this paradox together with the paradox of "unsignifiability" and explores the boundary between statements that are unproblematic in daily life and paradoxes.[4]

In early Islamic tradition liar paradox was discussed for at least five centuries starting from late 9th century apparently without being influenced by any other tradition. Naṣīr al-Dīn al-Ṭūsī could have been the first logician to identify the liar paradox as self-referential.[5]

Explanation and variants[edit]

The problem of the liar paradox is that it seems to show that common beliefs about truth and falsity actually lead to a contradiction. Sentences can be constructed that cannot consistently be assigned a truth value even though they are completely in accord with grammar and semantic rules.

The simplest version of the paradox is the sentence:

A: This statement(A) is false.

If (A) is true, then "This statement is false" is true. Therefore, (A) must be false. The hypothesis that (A) is true leads to the conclusion that (A) is false, a contradiction.

If (A) is false, then "This statement is false" is false. Therefore, (A) must be true. The hypothesis that (A) is false leads to the conclusion that (A) is true, another contradiction. Either way, (A) is both true and false, which is a paradox.

However, that the liar sentence can be shown to be true if it is false and false if it is true has led some to conclude that it is "neither true nor false".[6] This response to the paradox is, in effect, the rejection of the claim that every statement has to be either true or false, also known as the principle of bivalence, a concept related to the law of the excluded middle.

The proposal that the statement is neither true nor false has given rise to the following, strengthened version of the paradox:

This statement is not true. (B)

If (B) is neither true nor false, then it must be not true. Since this is what (B) itself states, it means that (B) must be true. Since initially (B) was not true and is now true, another paradox arises.

Another reaction to the paradox of (A) is to posit, as Graham Priest has, that the statement is both true and false. Nevertheless, even Priest's analysis is susceptible to the following version of the liar:

This statement is only false. (C)

If (C) is both true and false, then (C) is only false. But then, it is not true. Since initially (C) was true and is now not true, it is a paradox. However, it has been argued that by adopting a two-valued relational semantics (as opposed to functional semantics), the dialetheic approach can overcome this version of the Liar.[7]

There are also multi-sentence versions of the liar paradox. The following is the two-sentence version:

The following statement is true. (D1)

The preceding statement is false. (D2)

Assume (D1) is true. Then (D2) is true. This would mean that (D1) is false. Therefore, (D1) is both true and false.

Assume (D1) is false. Then (D2) is false. This would mean that (D1) is true. Thus (D1) is both true and false. Either way, (D1) is both true and false - the same paradox as (A) above.

The multi-sentence version of the liar paradox generalizes to any circular sequence of such statements (wherein the last statement asserts the truth/falsity of the first statement), provided there are an odd number of statements asserting the falsity of their successor; the following is a three-sentence version, with each statement asserting the falsity of its successor:

E2 is false. (E1)

E3 is false. (E2)

E1 is false. (E3)

Assume (E1) is true. Then (E2) is false, which means (E3) is true, and hence (E1) is false, leading to a contradiction.

Assume (E1) is false. Then (E2) is true, which means (E3) is false, and hence (E1) is true. Either way, (E1) is both true and false - the same paradox as with (A) and (D1).

There are many other variants, and many complements, possible. In normal sentence construction, the simplest version of the complement is the sentence:

This statement is true. (F)

If F is assumed to bear a truth value, then it presents the problem of determining the object of that value. But, a simpler version is possible, by assuming that the single word 'true' bears a truth value. The analogue to the paradox is to assume that the single word 'false' likewise bears a truth value, namely that it is false. This reveals that the paradox can be reduced to the mental act of assuming that the very idea of fallacy bears a truth value, namely that the very idea of fallacy is false: an act of misrepresentation. So, the symmetrical version of the paradox would be:

The following statement is false. (G1)

The preceding statement is false. (G2)

Possible resolutions[edit]

Alfred Tarski[edit]

Alfred Tarski diagnosed the paradox as arising only in languages that are "semantically closed", by which he meant a language in which it is possible for one sentence to predicate truth (or falsehood) of another sentence in the same language (or even of itself). To avoid self-contradiction, it is necessary when discussing truth values to envision levels of languages, each of which can predicate truth (or falsehood) only of languages at a lower level. So, when one sentence refers to the truth-value of another, it is semantically higher. The sentence referred to is part of the "object language", while the referring sentence is considered to be a part of a "meta-language" with respect to the object language. It is legitimate for sentences in "languages" higher on the semantic hierarchy to refer to sentences lower in the "language" hierarchy, but not the other way around. This prevents a system from becoming self-referential.

Unfortunately, this system is incomplete. One would like to be able to make statements such as "For every statement in level α of the hierarchy, there is a statement at level α+1 which asserts that the first statement is false." This is a true, meaningful statement about the hierarchy that Tarski defines, but it refers to statements at every level of the hierarchy, so it must be above every level of the hierarchy, and is therefore not possible within the hierarchy (although bounded versions of the sentence are possible).[citation needed]

Arthur Prior[edit]

Arthur Prior asserts that there is nothing paradoxical about the liar paradox. His claim (which he attributes to Charles Sanders Peirce and John Buridan) is that every statement includes an implicit assertion of its own truth. Thus, for example, the statement "It is true that two plus two equals four" contains no more information than the statement "two plus two equals four", because the phrase "it is true that..." is always implicitly there. And in the self-referential spirit of the Liar Paradox, the phrase "it is true that..." is equivalent to "this whole statement is true and ...".

Thus the following two statements are equivalent:

This statement is false.

This statement is true and this statement is false.

The latter is a simple contradiction of the form "A and not A", and hence is false. There is therefore no paradox because the claim that this two-conjunct Liar is false does not lead to a contradiction. Eugene Mills[8] and Neil Lefebvre and Melissa Schelein[9] present similar answers.

Saul Kripke[edit]

Saul Kripke argued that whether a sentence is paradoxical or not can depend upon contingent facts.[10][11] If the only thing Smith says about Jones is

A majority of what Jones says about me is false.

and Jones says only these three things about Smith:

Smith is a big spender.

Smith is soft on crime.

Everything Smith says about me is true.

If Smith really is a big spender but is not soft on crime, then both Smith's remark about Jones and Jones's last remark about Smith are paradoxical.

Kripke proposes a solution in the following manner. If a statement's truth value is ultimately tied up in some evaluable fact about the world, that statement is "grounded". If not, that statement is "ungrounded". Ungrounded statements do not have a truth value. Liar statements and liar-like statements are ungrounded, and therefore have no truth value.

Jon Barwise and John Etchemendy[edit]

Jon Barwise and John Etchemendy propose that the liar sentence (which they interpret as synonymous with the Strengthened Liar) is ambiguous. They base this conclusion on a distinction they make between a "denial" and a "negation". If the liar means, "It is not the case that this statement is true", then it is denying itself. If it means, "This statement is not true", then it is negating itself. They go on to argue, based on situation semantics, that the "denial liar" can be true without contradiction while the "negation liar" can be false without contradiction. Their 1987 book makes heavy use of non-well-founded set theory.[12]

Dialetheism[edit]

Graham Priest and other logicians, including J.C. Beall, and Bradley Armour-Garb have proposed that the liar sentence should be considered to be both true and false, a point of view known as dialetheism. Dialetheism is the view that there are true contradictions. Dialetheism raises its own problems. Chief among these is that since dialetheism recognizes the liar paradox, an intrinsic contradiction, as being true, it must discard the long-recognized principle of explosion, which asserts that any proposition can be deduced from a contradiction, unless the dialetheist is willing to accept trivialism - the view that all propositions are true. Since trivialism is an intuitively false view, dialetheists nearly always reject the explosion principle. Logics that reject it are called paraconsistent.

Non-cognitivism[edit]

Andrew Irvine has argued in favour of a non-cognitivist solution to the paradox, suggesting that some apparently well-formed sentences will turn out to be neither true nor false and that "formal criteria alone will inevitably prove insufficient" for resolving the paradox.[6]

Bhartrhari's perspectivism[edit]

The Indian grammarian-philosopher Bhartrhari (late fifth century AD) dealt with paradoxes such as the liar in a section of one of the chapters of his magnum opus the Vākyapadīya. Although chronologically he precedes all modern treatments of the problem of the liar paradox, it has only very recently become possible to confront his views and analyses with those of modern logicians and philosophers because sufficiently reliable editions and translations of his work have only started becoming available since the second half of the 20th century. Bhartrhari's solution fits into his general approach to language, thought and reality, which has been characterized by some as "relativistic", "non-committal" or "perspectivistic".[13] With regard to the liar paradox (sarvam mithyā bravīmi "everything I am saying is false") Bhartrhari identifies a hidden parameter which can change unproblematic situations in daily communication into a stubborn paradox. Bhartrhari's solution can be understood in terms of the solution proposed in 1992 by Julian Roberts: "Paradoxes consume themselves. But we can keep apart the warring sides of the contradiction by the simple expedient of temporal contextualisation: what is 'true' with respect to one point in time need not be so in another ... The overall force of the 'Austinian' argument is not merely that 'things change', but that rationality is essentially temporal in that we need time in order to reconcile and manage what would otherwise be mutually destructive states."[14] According to Robert's suggestion, it is the factor "time" which allows us to reconcile the separated "parts of the world" that play a crucial role in the solution of Barwise and Etchemendy.[15] The capacity of time to prevent a direct confrontation of the two "parts of the world" is here external to the "liar". In the light of Bhartrhari's analysis, however, the extension in time which separates two perspectives on the world or two "parts of the world" — the part before and the part after the function accomplishes its task — is inherent in any "function": also the function to signify which underlies each statement, including the "liar".[16] The unsolvable paradox — a situation in which we have either contradiction (virodha) or infinite regress (anavasthā) — arises, in case of the liar and other paradoxes such as the unsignifiability paradox (Bhartrhari's paradox), when abstraction is made from this function (vyāpāra) and its extension in time, by accepting a simultaneous, opposite function (apara vyāpāra) undoing the previous one.

Logical structure[edit]

For a better understanding of the liar paradox, it is useful to write it down in a more formal way. If "this statement is false" is denoted by A and its truth value is being sought, it is necessary to find a condition that restricts the choice of possible truth values of A. Because A is self-referential it is possible to give the condition by an equation.

If some statement, B, is assumed to be false, one writes, "B = false". The statement (C) that the statement B is false would be written as "C = "B = false"". Now, the liar paradox can be expressed as the statement A, that A is false:

"A = "A = false""

This is an equation from which the truth value of A = "this statement is false" could hopefully be obtained. In the boolean domain "A = false" is equivalent to "not A" and therefore the equation is not solvable. This is the motivation for reinterpretation of A. The simplest logical approach to make the equation solvable is the dialetheistic approach, in which case the solution is A being both "true" and "false". Other resolutions mostly include some modifications of the equation; Arthur Prior claims that the equation should be "A = 'A = false and A = true'" and therefore A is false. In computational verb logic, the liar paradox is extended to statements like, "I hear what he says; he says what I don't hear", where verb logic must be used to resolve the paradox.[17]

Applications[edit]

Gödel's first incompleteness theorem[edit]

Gödel's incompleteness theorems are two fundamental theorems of mathematical logic which state inherent limitations of sufficiently powerful axiomatic systems for mathematics. The theorems were proven by Kurt Gödel in 1931, and are important in the philosophy of mathematics. Roughly speaking, in proving the first incompleteness theorem, Gödel used a modified version of the liar paradox, replacing "this sentence is false" with "this sentence is not provable", called the "Gödel sentence G". Thus for a theory "T", "G" is true, but not provable in "T". The analysis of the truth and provability of "G" is a formalized version of the analysis of the truth of the liar sentence.[18]

To prove the first incompleteness theorem, Gödel represented statements by numbers. Then the theory at hand, which is assumed to prove certain facts about numbers, also proves facts about its own statements. Questions about the provability of statements are represented as questions about the properties of numbers, which would be decidable by the theory if it were complete. In these terms, the Gödel sentence states that no natural number exists with a certain, strange property. A number with this property would encode a proof of the inconsistency of the theory. If there were such a number then the theory would be inconsistent, contrary to the consistency hypothesis. So, under the assumption that the theory is consistent, there is no such number.

It is not possible to replace "not provable" with "false" in a Gödel sentence because the predicate "Q is the Gödel number of a false formula" cannot be represented as a formula of arithmetic. This result, known as Tarski's undefinability theorem, was discovered independently by Gödel (when he was working on the proof of the incompleteness theorem) and by Alfred Tarski.

George Boolos has since sketched an alternative proof of the first incompleteness theorem that uses Berry's paradox rather than the liar paradox to construct a true but unprovable formula.

In popular culture[edit]

As Argentinian writer Jorge Luis Borges comments in his prologue to the book Mathematics and Imagination,[19][20] an early version of the Liar's Paradox is formulated by Miguel de Cervantes in the 51th chapter of his book Don Quixote; where the following text can be read: "Sir, a deep river divides a certain lord’s estate into two parts… Listen carefully, your worship, for the case is an important one and rather difficult. I must tell you, then, that over this river is a bridge, and at one end a gallows and a sort of courthouse, in which four judges sit to administer the law imposed by the owner of the river, the bridge and the estate. It runs like this: “Before anyone crosses this bridge, he must first state on oath where he is going and for what purpose. If he swears truly, he may be allowed to pass; but if he tells a lie, he shall suffer death by hanging on the gallows there displayed, without any hope of mercy.” Though they know the law and its rigorous conditions, many people cross the bridge and, as they clearly make true statements the judges let them pass freely. Now it happened that they once put a man on his oath, and he swore that he was going to die on the gallows there—and that was all. After due deliberation the judges pronounced as follows: “If we let this man pass freely he will have sworn a false oath and, according to the law, he must die; but he swore that he was going to die on the gallows, and if we hang him that will be the truth, so by the same law he should go free."[21]

The liar paradox is occasionally used in fiction to shut down artificial intelligences, who are presented as being unable to process the sentence. In Star Trek: The Original Series episode "I, Mudd", the liar paradox is used by Captain Kirk and Harry Mudd to confuse and ultimately disable an android holding them captive. In the 1973 Doctor Who serial The Green Death, the Doctor temporarily stumps the insane computer BOSS by asking it "If I were to tell you that the next thing I say would be true, but the last thing I said was a lie, would you believe me?" However BOSS eventually decides the question is irrelevant and summons security.

In the 2011 videogame Portal 2, GLaDOS attempts to use the "this sentence is false" paradox to defeat the naïve artificial intelligence Wheatley, but, lacking the intelligence to realize the statement a paradox, he simply responds, "Um, true. I'll go with true. There, that was easy." and is unaffected, though other nearby AI in the room, intelligent enough to consider GLaDOS's statement, suffer a breakdown from hearing the paradox.

The second book in Emily Rodda's Deltora Quest series, The Lake of Tears has the main character, Lief, forced to answer a riddle correctly or be killed by the guardian of a bridge. When Lief answers the trick riddle wrongly, he confronts the guardian with his treachery. The guardian answers with another riddle, telling Lief to make a statement; if false, he will kill Lief by chopping off his head; if true, he will strangle Lief. Lief replies, "You will cut off my head." As the guardian was cursed to his fate by the evil sorceress Thaegan 'until truth and lies become one', the paradox allows him to revert to his original form: an eagle.

In the seventh episode of Minecraft: Story Mode titled "Access Denied" the main character Jesse and his friends are captured by a supercomputer named PAMA. After PAMA controls two of Jesse's friends, Jesse learns that PAMA stalls when processing and uses a paradox to confuse him and escape with his last friend. One of the paradoxes the player can make him say is the liar paradox.