Chapter 1

The Liar paradox

Suppose I say,

I am lying now.

Am I telling the truth, or am I lying? Suppose I'm telling the truth. Then what I say is the case – and so I'm lying. On the other hand, suppose I'm lying. But that's what I say. So I'm telling the truth. Either way, we are landed in contradiction: We are caught in the paradox of the Liar.

We get into the same kind of trouble if I say,

This sentence is false.

If what I say is true, then it's false; and if what I say is false, then it's true. We are confronted with the Liar again, under a slightly different guise.

Here is another version of the Liar, sometimes called the *heterological paradox*. Some words are true of themselves: For example, the word 'polysyllabic', itself a polysyllabic word, is true of itself; and so are the words 'significant', 'common', and 'prosaic'. Other words are false of themselves – the word 'new', for example, is not itself a new word, and so is false of itself; and so are the words 'useless', 'ambiguous', and 'long'. Let's call these words that are false of themselves *heterological*. But now take the word 'heterological' and ask whether *it* is true of itself or false of itself. If it is true of itself, then it is heterological, and so not true of itself; and if it is false of itself, then it is heterological, and therefore true of itself. And so we have a paradox.

Now these paradoxes arise out of quite ordinary concepts – in particular, the concepts of truth and falsity. The first two paradoxes arise directly from these concepts; in the case of the heterological paradox, we construct from our ordinary notion of falsity what seems to be a perfectly clear-cut semantic concept. There is nothing technical or recherché about the Liar paradox – it is quickly appreciated by ordinary speakers of English. This points to the significance of the Liar: It suggests that we do

Cambridge University Press
0521430690 - Universality and the Liar: An Essay on Truth and the Diagonal Argument - Keith Simmons
Excerpt
More information

not have a proper understanding of our ordinary notions of truth and falsity. An investigation of the Liar promises to correct and deepen our understanding of our basic semantic concepts. Then, and only then, will we be able to develop an adequate theory of truth and a satisfactory account of the relation between language and the world.

1.1. SOME VERSIONS OF THE LIAR

The Liar comes in many forms. The rich variety of Liar paradoxes is of intrinsic interest. But it has a further significance: the different versions of the Liar generate a constraint on any purported solution to the Liar. An adequate solution should deal with the Liar in *all* its manifestations, not just some. In my view, the Liar is as yet unsolved: no proposal has provided a satisfactory treatment of all the versions of the Liar we are about to see. Like any paradox, we should think of the Liar as a kind of argument, leading from apparently unexceptionable premises to an apparent contradiction, by apparently valid reasoning. Associated with distinct versions of the paradox are distinct Liar sentences. It is often convenient to identify a version of the paradox via its associated Liar sentence (or sentences).

Perhaps the simplest kind of Liar sentence involves just falsity and self-reference:

(1) This sentence is false.

Another simple Liar sentence involves truth, negation, and self-reference:

(2) This sentence is not true.

And, given the term 'heterological', we can construct this paradoxical sentence:

(3) 'Heterological' is heterological.

There are Liar sentences that display truth-functional complexity. Consider

(4) $2+2=4$ and this conjunct is not true.

If we suppose that (4) is true, then both conjuncts are true; but if the second conjunct is true, it follows that it is not true. And if the second conjunct is not true, then (4) is not true. On the other hand, if (4) is not true, then the second conjunct is true; and then, since the first conjunct is also true, (4) is true. We obtain a contradiction either way. Contrast (4) with

(5) $2+2\neq 4$ and this conjunct is not true.

Sentence (5) raises a new question: Should we evaluate (5) as false in virtue of its first conjunct, or should we evaluate (5) as paradoxical in virtue of its paradoxical second conjunct?

We may construct "Truth Tellers" as well as Liars. Consider

(6) (6) is true.

Unlike the Liar, we can assume that the Truth Teller is true, or false, without contradiction. (6) is not strictly paradoxical: no contradiction is generated from it. But the assignment of either truth value to the Truth Teller is entirely arbitrary. Even if not paradoxical, (6) is still semantically pathological. If we add a little truth-functional complexity to the Truth Teller, we find ourselves apparently able to prove any claim we like – we are landed in Curry's paradox.[1] Consider the following proof of the existence of God. We start with a version of the Truth Teller:

(7) If (7) is true then God exists.

We argue as follows: Assume

(a) (7) is true.

From (a) and the substitutivity of identicals, we obtain

(b) 'If (7) is true then God exists' is true.

Now truth is disquotational: To say a sentence is true is equivalent to asserting the sentence itself. So, in particular, we have the following biconditional:

(c) 'If (7) is true then God exists' is true iff if (7) is true then God exists.

From (b) and (c), it follows that

(d) If (7) is true then God exists.

From (a) and (d), we infer

(e) God exists.

Thus far, we have inferred (e) on the basis of (a) alone. So we may assert

(f) If (7) is true then God exists.

From (c) and (f), it follows that

(g) 'If (7) is true then God exists' is true.

From (g) and the substitutivity of identicals, we infer

(h) (7) is true.

And now, finally, from (d) and (h), we obtain our desired conclusion:

(i) God exists.

Thus far, we have considered only single sentences that generate semantic pathology. But Liar paradoxes may involve networks of sentences where the semantic status of one sentence depends on that of others. John Buridan's tenth sophism is:

> 10.0 *There are the same number of true and false propositions.*
> Let us posit that there are only four propositions: (1) 'God exists', (2) 'A man is an animal', (3) 'A horse is a goat', and (4) the above sophism. Given that situation, the question is whether the sophism is true or false.[2]

Given the truth of the first two propositions[3] and the falsity of the third, we arrive at a contradiction whether we assume the fourth is true or false.

Looped sentences provide another kind of paradoxical network. Consider this simple loop:

(8) (9) is true.
(9) (8) is false.

It's easily checked that we cannot consistently evaluate either sentence as true or as false. Here is a Truth Teller loop:

(10) (11) is true.
(11) (10) is true.

We can consistently evaluate both sentences as true, or both as false; but as with the Truth Teller, these evaluations are quite arbitrary. Loops can be of arbitrarily finite length:

(ρ_1) ρ_2 is true.
(ρ_2) ρ_3 is true.
⋮
(ρ_k) ρ_{k+1} is true.
(ρ_{k+1}) ρ_1 is not true.

There is no consistent assignment of truth values to the members of this loop. The upper limit on the length of a loop will be a delicate matter, turning on the cardinality of the language in which the sentences of the loop are expressed.

Besides loops, there are chains: infinite sequences of sentences without repetitions, where each sentence refers to the next. For example:

(σ_1) σ_2 is false.

(σ_2) σ_3 is true.
⋮
(σ_{2k-1}) σ_{2k} is false.
(σ_{2k}) σ_{2k+1} is true.
⋮

With this chain, there are consistent ways of assigning truth values.[4] Suppose, for example, we assign to σ_1 the value true; then we may assign the next pair of sentences σ_2 and σ_3 the value false, the pair σ_4 and σ_5 the value true, the pair σ_6 and σ_7 the value false, and continue to alternate in this way. Alternatively, we may assign to σ_1 the value false, to σ_2 and σ_3 the value true, to σ_4 and σ_5 the value false, and so on. Any such consistent assignment of truth values to the members of a chain is arbitrary.

Loops and chains may be intertwined. Consider this truth-functionally complex system of sentences:

(τ_1) τ_2 is false.
(τ_2) τ_1 is true and τ_3 is true.
(τ_3) τ_4 is false.
(τ_4) τ_3 is true and τ_5 is true.
⋮
(τ_{2k-1}) τ_{2k} is false.
(τ_{2k}) τ_{2k-1} is true and τ_{2k+1} is true.
⋮

There is no consistent way of assigning truth values to the sentences of this system.

Each version of the Liar we have seen so far has its "empirical" counterparts. For example, suppose you pass by my classroom and see that I've written something false on the board. Intending to expose your fraudulent colleague, you write

(12) The sentence written on the board in room 101 is false.

But you are yourself in room 101, and so the sentence you have written is a Liar sentence, an empirical counterpart of (1). Both (1) and (12) are Liar sentences; but (12) is a Liar sentence because the empirical circumstances so conspire. If the empirical circumstances had been different, if indeed I was in room 101, then your sentence would be straightforwardly true. Your sentence is not semantically ill-formed; it is not intrinsically pathological. Pathologicality here depends on the empirical circumstances. Parallel remarks can be made about the Truth Teller and Curry's paradox. An empirical version of (7) makes the "proof" of the existence of God even more striking.

Empirical loops are easily constructed. Plato might say,

(13) What Aristotle is saying now is true,

while, at that very moment, Aristotle is saying,

(14) What Plato is saying is false.

Here we have an empirical analogue of the pair (8) and (9). Again, there is nothing intrinsically ill-formed about (13) or (14): on other occasions, Plato or Aristotle may utter tokens of the same type that are straightforwardly true or false. Such empirical loops may be extended indefinitely, by adding participants in the appropriate way. And we can construct empirical chains, too. Suppose that each day someone goes to the Great Rock and says, "The next sentence uttered here will be true." A variant of this case will provide an empirical counterpart of the chain $\sigma_1, \sigma_2, \ldots$; and another variant will yield an empirical analogue of the system τ_1, $\tau_2, \ldots$.

We can produce more sophisticated empirical cases. Consider a case due to Saul Kripke.[5] Suppose Jones says,

(15) Most of Nixon's assertions about Watergate are false.

In ordinary circumstances, it would be straightforward to evaluate (15): we would list Nixon's utterances about Watergate and assess each for truth and falsity. But suppose that the empirical circumstances are such that Nixon's Watergate-related assertions are equally divided between the true and the false, except for this one:

(16) Everything Jones says about Watergate is true.

And suppose further that (15) is Jones's only utterance about Watergate (or that everything else Jones has said about Watergate is true). Then it is easily seen that, under these empirical circumstances, (15) and (16) are paradoxical.

We turn now to a version of the Liar that is sometimes called the *Strengthened Liar*. This is a version of the Liar that will occupy us quite a bit in later chapters. As we shall see, for some theories of truth the Strengthened Liar presents a stiff challenge; for other theories, it provides a starting point. Consider a simple Liar sentence:

(17) (17) is not true.

We may reason about (17) in the usual way and conclude that (17) is paradoxical. But if (17) is paradoxical, then it cannot be true. That is, we may infer

(18) (17) is not true.

But if we can truly assert (18), then (17) must be true after all, since (17) just says what (18) says.

Finally, we come to a type of Liar that is perhaps the most virulent. We may call it the *Revenge Liar*. There are different strains for different solutions to the Liar. If a solution to the Liar invokes truth-value gaps, so that a Liar sentence is treated as neither true nor false, then the Revenge Liar invokes these gaps too. Consider

(19) (19) is false or neither true nor false.

We are led to a contradiction whether we suppose (19) is true, false, or neither true nor false. In particular, invoking a truth-value gap will not help where (19) is concerned.

Similarly for other solutions: whatever semantic concepts are invoked by the solution, these form the basis for the Revenge Liar. If Liar sentences are diagnosed as semantically unstable, than a Revenge Liar is generated from

(20) (20) is not a semantically stable truth.

If we offer a Tarskian approach, according to which we split natural language into a hierarchy of object languages and metalanguages, each with its own distinct truth predicate, then we must deal with sentences like

(21) (21) is not true at any level of the hierarchy.

And a contextual theory, which avoids semantic paradox by the identification of shifts in the context of utterance, faces its own Revenge Liar:

(22) (22) is not true in any context of utterance.

Needless to say, some version of the Revenge Liar threatens to infect any attempt to produce a consistent theory of truth.

1.2. PROPOSALS

This wide range of cases should make us wary of taking too quick a way with the Liar. For example, it has been claimed that Liar sentences do not express propositions: Since Liar sentences do not really say anything, no paradox is forthcoming.[6] But more needs to be said. Empirical versions of the Liar create difficulties for undeveloped claims of this sort. Paul of Venice puts the point vividly: It is a consequence of the present claim that "some statement is not a proposition and becomes a proposition merely through a change in something else a thousand miles away."[7] Adapting Paul's example, suppose that someone says, "A false statement exists," and nothing else is said. Under these conditions, the statement is paradoxical. Now suppose that someone (a thousand miles away) makes a false statement. Then the original statement is no longer paradoxical, but true. It follows from the present claim that whether or not the original statement expresses a proposition depends on the empirical circumstances. But that shows that there is no intrinsic feature of the original sentence that determines whether or not it expresses a proposition. We lack an independent principle for distinguishing utterances that express propositions and those that don't. One cannot say that a sentence fails to express a proposition if it leads to semantic paradox, once all the empirical facts are determined. That would be plainly circular.

Some have suggested that Liar utterances fail to express a proposition because they are not syntactically well-formed.[8] In the face of empirical versions of the Liar, the proposal seems quite implausible. On such a view, (12) is not a syntactically well-formed sentence. However, under different empirical circumstances, a token of the same type as (12) is a perfectly well-formed sentence and does express a proposition. But how can syntactic well-formedness vary with the empirical circumstances?

Others have proposed that we ban self-reference: Liar sentences cannot express propositions because they are self-referential.[9] If we focus on simple Liar sentences, like (1) and (2), then the proposal seems natural enough: self-reference is a prominent feature of these sentences. But when we move on to loops and chains, then a simple ban on self-reference will no longer do. And the heterological paradox, generated by (3), does not involve any self-referential sentence.[10]

In this century, there have been two main kinds of response to the Liar (the two need not be mutually exclusive). One response is to reject classical logic and semantics. The other is a hierarchical response, associated with Russell and Tarski. These responses will be sketched now. In subsequent chapters, I shall argue against both responses and develop my own alternative.

1.2.1. Nonstandard logic and semantics

Nearly all who take this route abandon the *principle of bivalence,* according to which every sentence is true or false. Sometimes, a third truth value is introduced.[11] More usually in the literature, truth-value gaps are admitted.[12] According to the truth-value gap approach, paradoxical sentences are simply neither true nor false and have no other semantic value. Recall the version of the Liar generated by

(1) This sentence is false.

1.2 Proposals

The Liar reasoning depended on the assumption that (1) is either true or false. If we say that (1) is neither true nor false, but gappy, then the reasoning is blocked.

Proponents of this approach typically offer motivation for truth gaps independently of the Liar; invoking gaps just to get around the Liar would be unsatisfyingly ad hoc. Van Fraassen motivates truth-value gaps by appeal to Strawson's theory of presuppositions: The failure of a presupposition gives rise to truth-value gaps.[13] The notion of presupposition is characterized by

A presupposes B iff A is neither true nor false unless B is true.

For example, the sentence 'The present King of France is bald' presupposes the sentence 'The present King of France exists'. Since the later sentence is not true, the former is neither true nor false. According to van Fraassen, a Liar sentence like 'What I now say is false' "presupposes a contradiction, and hence cannot have a truth-value."[14]

Robert L. Martin's approach to the liar is motivated by category considerations: objects may fail to belong to the range of applicability of a predicate, so that certain predications yield "category-mistake sentences," sentences that are "semantically incorrect." Such sentences lack truth value. As examples, Martin cites "Virtue is triangular" and "The number 2 is green."[15] If we suppose that "heterological" does not belong to its own range of applicability, the contradiction of the heterological paradox is no longer derivable: The sentences '"Heterological" is heterological' and '"Heterological" is not heterological' are without truth value.[16]

Others have appealed to nonstandard semantics to give an account of vagueness.[17] The application of a vague predicate like 'bald' leads to sentences that are neither definitely true nor definitely untrue. It may be unsettled whether or not Harry is bald. But then it is likewise unsettled whether or not "Harry is bald" is true. According to Vann McGee, "The notion of truth inherits the vagueness of vague non-semantical terms."[18] This leads, according to McGee, to a tripartite division of sentences: those that are definitely true, those that are definitely untrue, and those that are unsettled. Liar sentences are unsettled.

There is another, more radical way of denying the principle of bivalence. We might allow sentences to be *both* true and false: there are truth-value "gluts," as opposed to gaps. A sentence like (1) is paradoxical, that is, both true and false. On such a *paraconsistent* approach, we embrace the contradiction generated by Liar sentences. This kind of approach has been taken by Nicholas Rescher and Robert Brandom, Graham Priest, and many others. We shall discuss paraconsistent accounts in Chapter 4.

1.2.2. Hierarchical views: Russell and Tarski

According to Russell, there is something that all the paradoxes have in common: Each involves an assertion that refers to an illegitimate totality. Consider the case of the Liar: "The liar says, 'everything that I assert is false'. This is, in fact, an assertion which he makes, but it refers to the totality of his assertions and it is only by including it in that totality that a paradox results."[19] Russell concludes:

> Whatever we suppose to be the totality of propositions, statements about this totality generate new propositions which, on pain of contradiction, must lie outside the totality. It is useless to enlarge the totality, for that equally enlarges the scope of statements about the totality. Hence there must be no totality of propositions, and "all propositions" must be a meaningless phrase.[20]

Similarly with the other paradoxes:

> All our contradictions have in common the assumption of a totality such that, if it were legitimate, it would at once be enlarged by new members defined in terms of itself.
>
> This leads us to the rule: "Whatever involves *all* of a collection must not be one of the collection," or, conversely, "If provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total."[21]

These are two formulations of Russell's vicious circle principle.[22]

As Russell goes on to say, this principle is negative in scope: "It suffices to show that many theories are wrong, but it does not show how the errors are to be rectified."[23] The theory of types is Russell's positive account of the paradoxes. Now Russell's ramified theory of types is notoriously complicated. For present purposes, we shall simplify considerably and restrict our attention to the hierarchy of orders of propositions.

Some propositions are themselves about propositions (e.g., the proposition 'Every proposition is true or false'); other propositions are not (e.g., the proposition 'Snow is white'). By the vicious circle principle, "Those that refer to some totality of propositions can never be members of that totality."[24] A hierarchy of orders of propositions is generated. "We may define first-order propositions as those referring to no totality of propositions; second-order propositions, as those referring to totalities of first-order propositions; and so on, *ad infinitum*."[25] And now the Liar is resolved: