Chapter 1

ASPECTS OF PARADOX

1.1. The liar and its variants

The liar is the paradox of the sentence that denies its own truth. In at least the usual versions of the paradox we have a sentence that either directly or indirectly denies its truth. In a typical form, the liar concerns the sentence (L):

(L) is not true.

We can easily end up in contradiction about that sentence.

One line of reasoning that leads to contradiction relies on the schema

(T) **S** is true iff *p*.

To get an instance of the schema, we must replace the letter '*p*' with a declarative sentence and the letter '**S**' with a name of that sentence. The name replacing '**S**' may either consist in the sentence itself put inside quotation marks or be different. However, I will call (T) 'disquotational' in the sense that, in each instance, an expression (sentence) is mentioned on the one side and used on the other. The schema (T) appears to be a principle that characterizes the concept of truth, since it captures the idea that a sentence is true iff things are as it says they are.

One instance of (T) is the biconditional

(L) is true iff (L) is not true.

Assume that (L) is true. Then, because of the biconditional, it is not true. Hence, by *reductio ad absurdum*, we can deny the assumption: (L) is not true. Consequently, because of the biconditional again, it is true—contradiction.

Instead of relying on (T), we can reach a contradiction by invoking two rules of inference: the rule that allows us to infer from a declarative sentence to calling it 'true' and the converse rule, which allows us to infer from calling such a sentence 'true' to the sentence itself. In classical logic those two rules, which we can name 'true-in' and 'true-out' respectively, are equivalent to (T). For, in classical logic, accepting the inference from a sentence **A** to a sentence **B** is equivalent to endorsing the conditional ⌜If **A** then **B**⌝. And, of course, accepting a rule of inference amounts to accepting all the inferences that conform with it, and endorsing a schema amounts to endorsing all its instances. So true-in is equivalent to the schema 'If p, then **S** is true', which captures the one direction of (T), and true-out is equivalent to the schema 'If **S** is true, then p', which captures the opposite direction. On the other hand, according to some other logical systems, accepting the inference from **A** to **B** does not by itself commit one to the conditional. Indeed, the logics presented in this book will be of that kind. Thus some theories of truth, which rely on such systems, endorse true-in and true-out but not (T). The theories developed in this book endorse both the rules and the schema.

The paradox has many versions. The sentence (L) is characterized by self-reference, but the paradox also arises without self-reference. In the sentences

(1) (2) is true
(2) (1) is not true,

we have circular reference and not self-reference, in that (1) contains a term denoting (2) and none denoting (1) while (2) contains a term denoting (1) and none denoting (2). By the schema (T), (1) is true iff (2) is true, and also (2) is true iff (1) is not true. Therefore, (1) is true iff it is not true, and the contradiction arises like before. Or again, take sentence

(3) There is a sentence that is written in line 25 of p. 1 and is not true.

Assume that the only sentence written in line 25 of p. 1 is (3). Then (there is a sentence that is written in line 25 of p. 1 and is not true) iff (3) is not true. By (T), (3) is true iff there is a sentence that is written in line 25 of p. 1 and is not true. Hence (3) is true iff it is not. But (3) does not refer to itself. It involves quantification over sentences, rather than reference to a sentence. The relation between (3) and itself is the relation between the sentence 'There is a person waiting at the stop' and Yannis Stephanou if he is waiting at the stop. Even if he is the only person waiting at the stop, the sentence does not refer to him, since it contains no name or other expression that denotes him.

(3) is one of the so-called *contingent* versions of the liar. In other words, things could be such that (3) was not involved in paradox, yet had the sense it also has actually. If one had not written (3) in line 25 of p. 1, but the sentence 'Naples is the capital of Italy', then (3) would not be involved in paradox; it would simply be true. It would also be true if both (3) and 'Naples is the capital of Italy' were written in line 25 next to each other. Of course, (L) and (1)–(2) are not contingent versions. A famous contingent example, due to Kripke [1984, pp. 54–55], consists of the sentences

(4) Everything Jones says about Watergate is true
(5) Most of the things that Nixon says about Watergate are untrue.

Let's imagine the following circumstances: Jones utters sentence (5) and says nothing else about Watergate; Nixon utters (4), and of the other things he says about Watergate, exactly half are true and the other half untrue. Then, by (T) and the circumstances, (4) is true iff (5) is true, and (5) is true iff (4) is not true. So the pair (4)–(5) becomes like (1)–(2). Yet in other circumstances no problem arises.

(4) and (5) could easily both be true, and they could easily both be false. They show that ordinary sentences about truth which in normal circumstances involve no contradiction can, in unusual but possible circumstances, become as paradoxical as (L). So the problem cannot be confined to a specifiable set of odd sentences, such as (L).

Some versions of the liar involve falsity rather than truth. Take the sentence

(6) (6) is false.

This sentence attributes falsity to itself. By (T), (6) is true iff (6) is false. This biconditional sounds paradoxical to many people, but in fact does not lead to contradiction unless supplemented with other principles about truth and falsity. One principle that seems reasonable is that if a sentence is false, then it is not true. So, by contraposition, if a sentence is true, it is not false. Assume that (6) is true; then by the biconditional it is false, and so, according to the principle just mentioned, it is not true. Hence, by *reductio*, the sentence is not true. Now assume that it is false; then by the biconditional it is true, so it is not false. Hence it is not false. (6) is neither true nor false. If we accept the principle of bivalence and believe that (6) is either true or false, we reach a contradiction. We may, however, reject bivalence and consider that (6) has no truth-value. Sometimes, the appellation 'liar' is restricted to a sentence like (6), and (L) is called 'a strengthened liar'. The idea is that in the case of (6) we can avoid contradiction by rejecting bivalence, but at least *prima facie* that option does not seem to help with (L).

In fact, rejection of bivalence does not seem to help with (6) either if we endorse the schema

(F) **S** is false iff not-p.

To get an instance of that schema, we must replace '**S**' with a name of a declarative sentence x and 'not-p' with a sentence that negates x. (F) captures the idea that a sentence is false iff things are not as it says they are. So it is as central to our understanding of falsity as (T) is to our understanding of truth. One instance of (F) is the biconditional '(6) is false iff (6) is not false', from which familiar steps lead to contradiction.

Also, if we adhere to (T) and accept that (6) is neither true nor false, we must be careful in our treatment of negation. If (6) is not false, then by (T)

($6'$) (6) is not false

is true. But ($6'$) negates (6). So we must allow for sentences that are neither true nor false, but whose negation is true. We should abandon the principle that a sentence not-**A** is true iff **A** is false. We may instead invoke another principle of standard semantics: not-**A** is true iff **A** is not true. At any rate, if we accept that (6) is neither true nor false, we must abandon one or other of those principles, whether or not we adhere to (T). Otherwise, we shall be committed to saying that ($6'$) is both true and not true: it is true because (6) is not true, and it is untrue because (6) is not false.

There are paradoxes about truth other than the liar. One of them is Curry's paradox [Curry, 1942]. Let the symbol '$\perp$' abbreviate any absurd sentence you want, e.g., the sentence '$3+1=5$', and think about this:

(C) If (C) is true, then $\perp$.

We can derive an absurd conclusion if we invoke both the schema (T) and the rule of conditional proof, that is, the rule that allows us, when we have made an assumption **A** and drawn a conclusion **B** within its scope, to infer (outside the scope of the assumption) the conditional ⌜If **A** then **B**⌝. By (T) we have:

(C) is true iff (if (C) is true, then $\perp$).

Assume that (C) is true. In that case, taking the left-to-right direction of the above biconditional, if (C) is true then $\perp$. Hence, by *modus ponens*, $\perp$. Therefore, by the rule of conditional proof, if (C) is true then $\perp$. Thus, now taking the right-to-left direction of the biconditional, (C) is true. Hence, by *modus ponens*, $\perp$. We have concluded that $3+1=5$. As Curry's paradox does not involve the concept of negation, it refutes the idea that the paradoxes arise from some problem in that concept.[1]

Another paradox about truth is Yablo's [Yablo, 1985, p. 340, and 1993], which involves the following sequence of sentences:

(Sentence 1)	For every $n > 1$, sentence n is not true
$\vdots$	$\vdots$
(Sentence m)	For every $n > m$, sentence n is not true
(Sentence $m+1$)	For every $n > (m+1)$, sentence n is not true
$\vdots$	$\vdots$

Assume that, for some number m, sentence m is true. Then, by (T), for every $n > m$, sentence n is not true; so, in particular, sentence $m+1$ is not true; hence, by (T) again, it is not the case that, for every $n > (m+1)$, sentence n is not true; then, there is a number $n > (m+1)$ such that sentence n is true, while it is also not true, which is absurd. Thus, by *reductio*, we may deny the assumption: for every number m, sentence m is not true. Therefore, sentence 1 is not true. But also, for every $n > 1$, sentence n is not true. Hence, by (T), sentence 1 is true—contradiction. In contrast with the liar, Yablo's paradox does not seem to involve any circularity.[2]

[1] One may consider that, for any sentence **A**, ⌜$\neg$**A**⌝ means ⌜If **A** then $\perp$⌝ and conclude that Curry's paradox involves the concept of negation. In fact, ⌜$\neg$**A**⌝ does not mean that. First, it may be defined as ⌜**A** $\to \perp$⌝, but such a definition does not capture the meaning of '$\neg$' any more than the definition of ⌜**A** $\wedge$ **B**⌝ as ⌜$\neg$[$\neg$**A** $\vee$ $\neg$**B**]⌝ captures the meaning of '$\wedge$'. Definitions of that kind are just a device for decreasing the number of primitive symbols in a logical system. Secondly, '$\perp$' may mean '$3+1=5$', but ⌜$\neg$**A**⌝ does not mean ⌜If **A** then $3+1=5$⌝. Why should it mean that and not ⌜If **A** then $2=1$⌝? And if we do not give the symbol '$\perp$' any particular sentential meaning, we have given no meaning to the sentence ⌜If **A** then $\perp$⌝ and so there is no candidate for synonymy with ⌜$\neg$**A**⌝.

[2] Priest [1997] began a discussion about whether Yablo's paradox really involves no circularity. [Cook, 2014] is a book-length study of the paradox.

Yablo's paradox, Curry's and the liar are all semantic paradoxes, since the concept of truth is a semantic notion. There are paradoxes that employ other semantic notions, such as reference, satisfaction, and so on. One of them, akin to the liar, is Grelling's paradox [Grelling and Nelson, 1908, p. 307]. Let us stipulate that a predicate is *heterological* iff it does not satisfy itself; in other words, it is heterological just in case it is not true of itself. So 'long' is heterological, but 'short' is not; for 'long' is not a long predicate, but 'short' is a short one. At least for monadic predicates, the concept of satisfaction seems to be characterized by the following schema:

(S) For everything x, x satisfies **G** iff x is F,

where the letter 'F' is to be replaced by such a predicate while '**G**' is to be replaced by a name of that predicate. For instance, anything satisfies 'long' iff it is long, and anything satisfies the predicate 'heterological' iff it is heterological. Thus if 'heterological' is heterological, then by the definition of 'heterological' it does not satisfy itself, and so by (S) it is not heterological. If, on the other hand, it is not heterological, then by the definition it satisfies itself, and so it is heterological. The predicate is heterological just in case it is not.

A satisfactory treatment of the liar should tell us, on the one hand, where the blame lies for the problem and, on the other, how we can overcome it. In other words, it should both offer a diagnosis and suggest a therapy.[3] Also, it ought to be able to deal with all the versions of the paradox and, if possible, with the other semantic paradoxes as well. Setting aside the approach which endorses contradictions and considers that the sentence (L) and the other paradoxical sentences both are and are not true, we can say that a satisfactory treatment should explain, on the one hand, where the blame lies for the production of contradictions and, on the other, how we can avoid them. We end up in contradictions relying on the schema (T) and using classical logic. (In the contingent versions, we also presuppose some facts that cannot be doubted and are to do with who says what, what is written where, and the like.) Much of the work that has been done on the paradox develops, in various ways, the idea that the blame lies in (T). My own work is included in the approach on which the blame lies in classical logic.

1.2. Propositions and truth-values

Those who consider that the blame for the problem lies in (T) do not, of course, need to reject the schema entirely. They can argue that we ought to restrict it: refuse to accept its instances that concern paradoxical sentences, but accept its instances that concern other sentences. Paradoxical sentences are just those in whose case we are led to contradiction. (L), (1) and (2) are paradoxical. (4) and (5), too, would be paradoxical if the appropriate circumstances obtained. That

[3] The distinction between the two tasks is essentially that made in [Chihara, 1979, pp. 590–591].

approach to (T) is particularly sensible if there are reasons to exclude paradoxical sentences from (T) which are independent of the fact that, in the case of those sentences, we are led to contradiction. But are there any such reasons?

1.2.1. No proposition expressed. One may argue that we should not apply (T) to sentences that express no proposition. In other words, we should not apply it to sentences that express no information about how things are.[4] The schema (T) is a formulation of the idea that a sentence is true iff things are as the sentence says they are. Obviously, the idea concerns only sentences that express information about how things are. Correspondingly, schema (T) should be accepted only to the extent that it concerns such sentences. But paradoxical sentences, one may continue, do not express propositions, do not convey information. Hence, we should not apply (T) to paradoxical sentences.

Indeed, one may offer the following argument in order to show that paradoxical sentences do not express propositions. It is not the case that (the sentence (L) is true iff it is not true). For if we have any biconditional of the form '*p* iff not-*p*', its negation is an instance of a logical law. Likewise, either it is not the case that (sentence (1) is true iff (2) is true) or it is not the case that (sentence (2) is true iff (1) is not true). Also, if the only sentence written in line 25 of p. 1 is (3), we must deny that (3) is true just in case there is a sentence that is written in line 25 of p. 1 and is not true. Other paradoxical sentences are similar. But then, paradoxical sentences have no truth-conditions. For if, e.g., (3) has a truth-condition, the condition can only be that there is a sentence that is written in line 25 of p. 1 and is not true. If paradoxical sentences have no truth-conditions, then they do not express propositions. A sentence expressing a proposition conveys information about how things are and so has a truth-condition.[5]

In my opinion, paradoxical sentences express propositions, although the opposite view is quite common; see, e.g., [Kripke, 1984, pp. 63–64]. The argument in the preceding paragraph ignores the possibility that classical logic may not apply to paradoxical sentences. Let's take (L). It is a negation of the following sentence:

---

[4] A *proposition* is such a piece of information. The information may be correct, even tautological; but also it may be wrong, even absurd. Declarative sentences, at least normally, express propositions. For example, the proposition that Kant is the most important German philosopher is expressed by the sentence 'Kant is the most important German philosopher' and every synonymous sentence of either English or another language. Interrogative and imperative sentences do not express propositions.

[5] One may claim that it is a misuse of 'true' to call a sentence 'true': only propositions can be appropriately so called, so all the sentences we have discussed present a category mistake. But one gains nothing by claiming that. For if only propositions can be appropriately called 'true', we can replace (L) and (T) with

(L^P) (L^P) does not express a true proposition

and

(T^P) **S** expresses a true proposition iff *p*.

We can similarly modify the other sentences we have discussed. Then, we shall face the various versions of the paradox again.

(L′) (L) is true.

If (L′) expresses a proposition, a piece of information, then (L) also expresses a proposition, the negation of the one expressed by (L′). It is clear what (L′) refers to: it refers to a certain series of words which is syntactically structured in a particular way and so is a sentence. It is clear what it says about that series: that it is a true sentence. So (L′) is an attribution of a property to a linguistic entity. How can it fail to express information about that entity?

The idea that paradoxical sentences do not express propositions ceases to be attractive as soon as we realize that there are propositions which are similar to paradoxical sentences and are paradoxical themselves. In order to reach a contradiction that concerns such propositions, we need the schema

(TP) *P* is true iff *p*.

We get an instance of (TP) when we replace the letter '*P*' with a term that refers to a proposition and the letter '*p*' with a sentence that expresses just that proposition. For example,

> The proposition expressed by the sentence 'Snow is white' is true iff snow is white.

The schema (TP) is the analogue of (T) for propositions.

We can see the sentence

(7) The proposition expressed by (7) is untrue.

Does (7) express a proposition? Yes, if there is a proposition to the effect that the proposition expressed by (7) is untrue. Is there such a proposition? Yes, at least according to the following argument from [Horwich, 1998, p. 41]: For any condition *C*, there could be someone believing that the proposition that satisfies *C* is untrue. Whatever could be believed by someone is a proposition. Thus, for any condition *C*, there is a proposition to the effect that the proposition that satisfies *C* is untrue. Hence there is a proposition to the effect that the proposition expressed by (7) is untrue. Moreover, there cannot be two propositions to that effect, so (7) expresses just one proposition.

It may be objected that unless we have a deflationary concept of a proposition, we should not accept that whatever could be believed by someone is a proposition; someone could believe that he was in tune with the universe, but surely there is no such proposition in any substantial sense of the term. My concept of a proposition is just the concept of a piece of information; it is not more substantial than that. Even if one doubts that whatever could be believed is a proposition, one should not doubt that if someone may tell someone else that things are a certain way, then there is a piece of information, and so a proposition, to the effect that things are that way. To tell someone that *p* is to convey information. And we may be told that the proposition expressed by (7) is untrue.

Let's call the proposition expressed by (7) 'Π'. According to (TP), Π is true iff the proposition expressed by (7) is untrue. Therefore, Π is true iff it is untrue. Π is similar to the sentence (L). Just like (L), it refers to itself and denies that it is true. The only difference is that (L) refers to itself by means of a name, whereas Π refers to itself by means of a description; it refers to itself as the proposition expressed by a certain sentence. (On the other hand, (7) refers to itself by means of a name.)

We can also see the sentence

(8) There is a proposition that is expressed by (8) and is not true.

Does (8) express a proposition? Yes, as we can see by comparing (8) to the sentence

(9) Some proposition is untrue and is expressed by (8).

(8) and (9) are distinct sentences, since they are not made up of the same words. (9) does not refer to itself, and this may make it easier to study. It seems clear to me that (9) expresses one (and only one) proposition. It is the information that some proposition satisfies two particular conditions (it is untrue, and it is expressed in a certain sentence). But sentences (8) and (9) are synonymous; they have no difference in sense. The one results from the other through small changes in wording alone. Therefore, (8), too, expresses a proposition, just the one expressed by (9). (Incidentally, here is a moral: self-reference is not an aspect of the sense of a self-referring sentence, since it is possible for another sentence to have just the same sense without being self-referring.)

Let's call the (only) proposition expressed by (8) 'Σ'. Thus (there is a proposition that is expressed by (8) and is not true) iff Σ is not true. By (TP), Σ is true iff (there is a proposition that is expressed by (8) and is not true). Hence, Σ is true iff it is not. Σ is similar to sentence (3). Σ says that there is a proposition which, on the one hand, satisfies a certain condition and, on the other, is not true; and Σ itself is the only proposition that satisfies the condition in question. Correspondingly, (3) says that there is a sentence which, on the one hand, satisfies a particular condition and, on the other, is not true; and (3) itself is the only sentence satisfying that condition. The similarity gets up to the contingency in the satisfaction of the relevant condition: just as (3) might not be written in line 25 of p. 1, so Σ might not be expressed by (8), since (8) could have an entirely different sense from the one it actually has and so fail to express Σ.

The proposition expressed by the liar sentence (L) is also paradoxical. Let's call it 'Λ'. Λ is not a self-referring proposition, since it refers to a certain sentence and not to itself or any other proposition. In that respect, it is less similar to (L) than Π is. By the schema (TP) we have

Λ is true iff (L) is not true.

In order to reach a contradiction about Λ, we can invoke the principle

> For every sentence S and every proposition P, if P is the proposition expressed by S, then S is true iff P is true.

By that principle, (L) is true iff Λ is true, and so (L) is not true iff Λ is not true. Hence, Λ is true iff it is not.

When one realizes that there are propositions like Π and Σ, one has no motivation any more to assert that sentences such as (L) and (3) express no proposition. The idea that paradoxical sentences express no proposition is appealing when one hopes that it will lead to a principled treatment of the liar and kindred paradoxes. But we cannot transpose the idea into the case of paradoxical propositions, so we cannot use it to tackle the problem in their case, so we cannot use it to tackle the problem in general.[6]

1.2.2. Lack of truth-value. Varying the idea that paradoxical sentences express no proposition, one may invoke another reason to exclude them from schema (T) which is independent of the fact that, in the case of those sentences, we are led to contradictions. One may argue that we should not apply (T) to any sentence that has no truth-value. For if we do, then the right hand-side of the resulting biconditional will be the sentence itself, so it will be neither true nor false. But the left-hand side, which describes the sentence as being true, will be just wrong, that is, false. And then, according to the usual systems of three-valued logic, the biconditional will not be true; it will be neither true nor false. Paradoxical sentences, one may continue, have no truth-value. Hence we should not apply (T) to such a sentence.

But why say that paradoxical sentences have no truth-value? Here it may be argued that paradoxical sentences are *ungrounded* and that ungrounded sentences, whether paradoxical or not, lack truth-value.[7] Sentence

(10) Snow is white

is not about truth or falsity, and its truth-value has consequences for other sentences. So

(10′) (10) is true

is true because (10) is true. The truth-value of (10) determines a truth-value for (10′). It also determines a truth-value for

(10″) (10′) is true
(10‴) (10″) is true
⋮ ⋮

as well as

(10+) Every sentence named by the numeral ‘(10)’ with or without primes is true.

Sentences (10′), (10″), . . . , (10+) are all true because (10) is true. A sentence *s* is grounded provided either it has a truth-value without being about truth or falsity (or other semantic issues) or there are some other sentences (one or more) which

[6] Burgess and Burgess [2011, p. 119] argue that paradoxical sentences can be used in explanation of action and so should express propositions.

[7] Herzberger [1970, pp. 147–151] first introduced a notion of groundedness. The explanation provided here is different from his.

are not about truth or falsity and whose truth-values determine a truth-value for *s*. (10), (10′), (10″), . . . , (10+) are grounded. If the sentence 'Naples is the capital of Italy' were written in line 25 of p. 1, then the truth-value of that sentence, together with the fact that it is written in a certain line, would determine a truth-value for (3); this would be sufficient for (3) to count as grounded.

On the other hand, let's see sentence

(11) (11) is true.

(11) is called a *truth-teller*. It is not paradoxical. Neither the assumption that it is true nor the assumption that it is false leads to contradiction. But it is ungrounded; there are no sentences that are not about truth or falsity and determine a truth-value for (11). Let's also see sentences

(12) (13) is true
(13) (14) is true
(14) (12) is true.

They are not paradoxical either, but they are ungrounded, since none of them has a truth-value due to sentences that are not about truth or falsity (or other semantic issues). Whether (12) is true depends on whether (13) is true, and on whether (14) is true, but it depends on no sentence that does not involve truth. Things are similar if instead of a circle we have a descending sequence:

(15) (15′) is true
(15′) (15″) is true
⋮ ⋮

So it may be argued that if a sentence possesses a truth-value, then either the sentence is about the real world, as it were, or at least its truth-value is due to the truth-values of some sentences that are about the real world. And it is supposed that the real world is not to do with whether sentence so-and-so is true, or whether sentence so-and-so is false, but with whether snow is white, whether there are extraterrestrials, and the like. If so, ungrounded sentences possess no truth-value.

Indeed, paradoxical sentences are ungrounded. Whether (L) is true depends on whether it itself is not true, but it depends on no sentence that does not involve truth or falsity. Whether (1) is true depends on whether (2) is true, and thus on whether (1) itself is not true, but it depends on no statement that is not about truth. I know of no variant of the liar that concerns a grounded sentence. However, it is difficult to cling to the view that all ungrounded sentences lack truth-value. Take the ungrounded sentence

(16) If (16) is true, then (16) is true.

It has the form 'if *p* then *p*'. Sentences of that form are entirely tautological, so we should not doubt that they are true. (16) is, therefore, true. Or take the ungrounded sentence

(17) (17) is true and (17) is not true.