Part III

PHILOSOPHICAL DIMENSIONS OF LOGICAL PARADOXES

7

Logical Paradoxes

JAMES CARGILE

Logical paradoxes centrally involve difficulties in determining truth values. But not all such difficulties are paradoxes and not all paradoxes are paradoxes of logic. Considerable trouble can be taken in trying to delineate the right subclass. But the questions about truth value are often more interesting. It is better to begin by trying to answer some notable problems of this sort, even at the risk of contributing to some subject other than logic.

One such case, is the Eubulidean Liar (UL), attributed to Eubulides of Megara, who is supposed to have said "What I am saying is false." This could be an unproblematic assertion made by a spy to an assistant as an aside about some item of misinformation he is in the process of sending out in a broadcast. But when understood with a certain kind of 'self-reference' the remark would be absurd. It was never offered as a sincere effort at communication, but as a way of presenting a problem for rationalistic philosophy. This problem may be better understood in the following version. The sentence

(A) The sentence A is not true,

is one which it seems could not be true. For if it were true, it would seem to follow that it is not true. But if we conclude from this that the sentence A is not true, then it seems this could not be right, being the very same words as the sentence A itself, which would seem to suggest that A is true after all. This has led to the suggestion that allowing truth to be attributed self-referentially, as in 'No proposition is both true and false' should be somehow avoided or restricted.

The most reasonable way to follow that idea would be to deny that there is any such property as truth. There is obviously an English predicate '. . . is true' and the grammar of A is unassailable. We can grammatically assert the sentence and grammatically attribute to it its predicate. But this is no guarantee of asserting a proposition and attributing a property. Formal logic is primarily concerned with sentences and predicates. But philosophical logic must be concerned with propositions and properties. Both propositions and sentences can be asserted or said and both predicates and properties can be predicated or attributed or said of. But propositions and properties are more important for philosophy, which makes for problems in the assimilation of formal logic and its impressive results. The great precision and secure consistency of some systems of formal logic may seem to point in favor of nominalism. But we can continue to value this precision while keeping open the possibility of a consistent use of propositions and properties in philosophy.

The primary connection between asserting or saying, and predicating or saying of, is the property of truth. We have Rule R1: To assert a proposition is one and the same thing as to predicate truth of it, and to deny a proposition is the same as predicating nontruth, which in application to a proposition is the same thing as falsity. To call a thing a nontruth is not to call it false. But to call it a nontrue proposition is to call it false, and that is how it is natural to understand calling a proposition nontrue – as merely short for 'nontrue proposition.' This fundamental connection does not apply at all in the case of believing. To believe that a proposition is true is not the same as believing it. One may believe that what Bill will say tomorrow is true without believing what Bill will say tomorrow, but to assert that what Bill will say tomorrow is true is to assert that thing, whether or not you know what proposition it is. To assert a proposition is to take a certain unique kind of responsibility for its being true. This is often done knowingly, but that an agent asserts a proposition does not in general entail that the agent believes it or even has the ability to understand it or know what proposition it is.

The ruling that A does not express a proposition seems to be good reason to conclude it is not true. But then it seems that is a true judgment about A which is expressed by A. To avoid that problem it would be tempting to conclude there is no such property as truth. A better response is to assume that A expresses some proposition. Whatever it is, A says it. But A says that whatever it says is not true. So by the above R1, A both says and denies the same proposition, saying whatever it says and that whatever it says is not true. This gives us an adequate basis for saying that A is not true.

It will be objected that this is just what A says. On the contrary, we cannot adequately say in full what it is that A says. It is not that what A says is that A is not true and also that what A says is that it is not true that A is not true. It is just that these are equally good representations of what A says. That is good enough to show that whatever A says is contradictory. It is not required that those equally good representations, which are as good as we can get, are good enough to warrant either one being taken as making it clear to us what A says. A is one of many counterexamples to the idea that to identify the proposition expressed by a sentence S it suffices to write 'the proposition that . . .' followed by the sentence S. Thus it would be quite wrong to think that 'The sentence A is true as a sentence of English if and only if the sentence A is not true' is licensed by a correct rule for describing the content of a sentence of English.

This may be reinforced by considering

(A′) The sentence A′ is true.

We can say that what the sentence A′ says is that what it says is true. That would indeed be, by R1, to say whatever it is that A′ says. That does not tell what A′ says or offer any ready guide as to what its truth value would have to be. If we rule that A′ says nothing, then we should treat anyone who claims that A′ is true as speaking falsely. He might have used the same words as A′. So why not count A′ as false also?

If a man says 'What I am now saying is true' (when it is clear there is no other reference) then he cannot be serious and we are right to rule that he has not asserted any proposition. Having so ruled, we must hold that one who says 'What he said then was true' (when the reference is clear) is speaking falsely. This can be explained by appeal to the thoughts expressed. When we treat sentences by themselves as doing the saying, matters cannot be clarified by that means. It might be suggested that we should not treat sentences in this way, but that is not a practical possibility. It is often important to ask, not what the author of certain words, such as a constitution, intended, but what has been said by them. Whether the founders of the USA intended their words to be incompatible with the institution of slavery or not, it is important that the words were not compatible with it.

Denying that A or A′ say anything seems to provide good reason to call A true and A′ false. This then leads to paradox. We do better to note that if A says anything, that thing is contradictory and thus false. Thus prepared, it is best to rule that A says something, albeit an obscure and worthless thing. If A′ says anything, we have no reason whatever to consider this thing to be false. To the extent that we have no reason either to consider it true, we must consider this a bad mark against the practice of taking sentences by themselves as saying things at all. But the consideration that A′ is doing absolutely nothing but endorsing whatever it is that it says may suffice as a reason for counting it trivially true. For every saying endorses itself, being equivalent to calling itself true. A′ may be taken to report this triviality about itself.

The Epimenidean Liar (EL) can be put as follows: we build a one room shed known as Building B, working in total silence. One of us then goes in and asserts

(B) Nothing true is asserted in Building B at any time

and nothing else. We then burn Building B to ashes. It seems that B cannot be true, since it was asserted in Building B and having something true asserted there would make B false. But if it is not true that no truth is asserted in B then it seems to follow (since something has been asserted) that some truth has been asserted in B. Since this cannot be B, and given the history of the building, we might then have to conclude that elves or similar beings slipped in and made at least one false assertion while the building was there to house such assertions. But this is hard to bear. We seem to have to avoid contradiction only by accepting a preposterous factual claim.

Here we appeal to the principle R2: that to assert that all Xs are Ys is to predicate being a Y of every X. Thus to assert B in the building is to predicate nontruth of everything asserted in the building. That is to assert that it is not true that everything asserted in B is nontrue in the very course of asserting that everything so asserted is nontrue. So the assertion of B in the building is false. Our assertion out of the building is more fortunate, since it is not among the assertions it is calling nontrue.

It would not be clear to say that our assertion of sentence B is not self-referential while the in-building assertion of it is, due to ambiguity concerning 'self-reference.' All assertions are self-referential in the sense that to assert any proposition P is to assert that everything whatsoever (including P) is such that P. (That (x)P is equivalent to P has been questioned for sentences of predicate logic in the 'empty domain' case, but this is an extremely eccentric system which should not influence our considerations.) In that sense, both the in-building assertion of B and our assertion of it are self-referential. On another interpretation of 'reference' we do not say that 'All Fs are Gs' is about itself unless it is an F. The in-building assertion of B says that if it is an in-building assertion then it is not true. In that sense it is self-referential, while our judgment about B is not an in-building assertion, so that in this other sense of 'reference' it is not self-referring.

The present line, that sentences such as B serve to reject all assertions of a given kind, has been advocated by some thinkers who go further, to hold that such sentences do not convey any additional assertion beyond each of those denials. Similarly, it is said that to say 'Everything he says is true' is merely to endorse everything he says and not to say anything further. This idea has the consequence that there would be no difference between asserting B in the building or outside it. Similarly, if someone asserts 'Every Cretan assertion is false' it would be irrelevant to the assessment of this performance whether it was put forward by a Cretan. This is quite implausible. The paradoxes are best answered, not by economies about meaning, but by paying attention to the full meaning.

We may contrast problem cases superficially similar but involving belief rather than assertion. Suppose that we build a Building C on exactly the lines of B, except that the only person to go in does not know what building he is in. He believes (never mind how) that Building C has been so constructed that any carefully considered belief held therein is false. While waiting there to be called upon to leave, he reflectively believes (as opposed to his unarticulated assumptions that he is in a building, clothed, etc.) only that

(C) Any beliefs reflectively held in Building C are false.

If that were the only such belief in room C then we would appear to have a situation similar to the EL. But the principle R2 is obviously false for beliefs. To believe that all C beliefs are false is not to believe that it is false that all C beliefs are false. However, to believe that all C beliefs are false (as opposed to merely believing such a thing as that the sentence C expresses a truth) is to believe that you are not in building C. This is not a separate belief, but rather, part of what it is to believe that all C beliefs are false. To assert that all C beliefs are false is to predicate the falsity of every belief reflectively held in Building C. The sincerity of that assertion would mean believing, not everything that is thereby asserted, but only that all C beliefs are false. The assertor would unknowingly predicate falsity of the belief he expresses in making the assertion. An assertor could be fully informed as to what he is doing in asserting C, but then in order to persist he must be insincere. To believe that all C beliefs are false is to believe among other things that that belief is not one reflectively held in Building C. It is not possible to believe *directly* (a notion which cannot be clarified further here) that your very belief is false (or that it is true). There is no belief analogue to the Eubulidean Liar. The indirect cases, such as C, always involve more content in the belief than is assumed in the formulation of apparent conflicts about truth value.

Two innocent but logically acute persons might be conversing about C, one of them standing in the yard of Building C, the other speaking to him from inside that building, neither one knowing that building to be C. They could both believe that all C beliefs are false and have essentially the same belief. It would be a belief including the mistaken thought that the building housing a party to their conversation is of course not the

Building C which they are conversing about. But if they asserted their common belief, they would, unbeknownst to either, make different assertions.

Suppose that Bill declares that P, Q, R, S, and T, and Bob and Ben hear his declaration and agree that everything that Bill declared is true. Bob remembers well that Bill said that P, Q, R, S, and T, while Ben is unable to recall just what Bill said. Here we could have many degrees between having only a confidence in Bill and not knowing at all what he said, perhaps even disagreeing with those propositions while ignorant that they are what Bill said, to being in Bob's state of complete understanding. To say that Bob and Ben agree on the proposition that everything that Bill said is true badly underdescribes this situation. We might best describe Bob as believing that Bill's declaration was that P, Q, R, S, and T and that as a matter of fact, P, Q, R, S, and T. Bob does agree with Ben that 'everything Bill said is true.' But what this belief consists in would need to be worked out in dialogue between them, so that Bob's belief is reduced to Ben's or Ben's expanded to Bob's or to some intermediate compromise.

It is only in successful dialogue that we achieve a good understanding of what is believed. A proposition is essentially something which can be conveyed to others in successful dialogue, or a compound of such things. (It is the compounding that allows for unbelievable propositions of various kinds.) Logic generalizes about these things, and in thus stepping back from specific dialogue, loses track of the identity of the propositions and treats merely of sentences. In extreme cases we have sentences such as A or A′, which could never be used, in their logically problematic roles, in good dialogue, as expressing contributions to the dialogue, though they can of course be objects of dialectical discussion.

Besides the belief cases, there are puzzles in which someone fears that all his fears are unfounded or hopes that all his hopes are unfulfilled, etc. Believing, hoping, fearing, and asserting are all things done by people, but the former are attitudes, while asserting is not. Finding common 'propositional objects' and restricting 'self-reference' or rejecting a 'global truth predicate' or employing evaluation rules which do not assign truth values to all propositions makes possible a uniform treatment of these puzzles. But this is a costly uniformity which blurs important differences. It is important to note the distinction between cases which require talking with someone who either expresses an attitude or attributes one to someone else, and cases which involve just looking at the powers of a sentence by itself.

For example, when someone claims an odd belief, we need to talk with that person rather than making adjustments in logic. If a man claims to fear that all his fears are unfounded we need to know if he intends to express fear that, among other things, his fear that all his fears are unfounded is unfounded. If he does, then the problem is not for logic, but for those who think this could be sincere.

The Geach–Lob implication liar (IL) involves

(D) D materially implies that P.

It seems that if D were false, then it would have to be true (and P false). Since that is impossible, it seems that D is necessarily true, and thus that P is too. A suitable choice of P can bring out how bad this would be. This paradox for the material conditional is not essentially different from other paradoxes based on other truth functional connectives. Here we appeal to the rule R3: that to assert that if P then Q and that P, in one assertion, is to assert that Q. Now the suitable choice of Q only yields a bad assertion, not a bad problem. When P = (2 + 2 = 4) that version of D is true. When P = (2 + 2 = 5) that version is false. Similar considerations apply to cases based on other connectives. (Our earlier case, A, could have been interpreted disjunctively, as saying that A either expresses no proposition, or expresses a false proposition, and that would require a principle for disjunction.)

The Grelling Liar (GL) involves the predicate 'heterological' defined as "a predicate which expresses [as a term in some systematic usage – expression cannot be analyzed here] a property of which it is not itself an instance." This self-reference seems to threaten both the saying that 'heterological' is heterological and the saying that it is not. The answer is that there is no such property as that of being a predicate which expresses a property of which it is not itself an instance, any more than there is such a property as being a property which is not an instance of itself. It is common for logicians to agree with this. But the crucial problem is to properly explain why it is so.

It is not that to say a term is heterological is to say nothing of it. It is rather, that it is not to say the same thing for every term. 'Heterological' expresses a property in application to 'obscene,' but the property is that of expressing the property of being obscene while not possessing it. In application to 'English' the property is that of expressing the property of being English while not possessing it. Even this much is often accepted. The question still remains as to why it is so. For one may of course assert, about a term whose meaning is unknown, that it is heterological. Why is this not merely to say that there is some property it expresses but does not possess? (It is widely held to be an important point of logic that to say that some F is a G is *not* to predicate being a G of any F.)

It is because R4: to assert that Some F is a G is to assert that if anything whatever is such that everything other than it is not an F which is a G, then that thing is an F which is a G. When we say that there is some property expressed by 'obscene' which is not possessed by it, we assert of each thing there is that if nothing other than it is expressed by 'obscene' then it is expressed by 'obscene' and not possessed by it. The one and only thing which satisfies the antecedent condition of this conditional predication is the property of being obscene, and so, for that reason, calling 'obscene' heterological is to predicate not being obscene of it. By contrast, the predicate 'heterological' fails to *uniformly* express any property. It always picks up its property from the term to which it is applied, so that when applied to itself, there is no property to pick up. For that reason it is not heterological – it does not possess a property it expresses, because it does not express a property. (If we define 'heterological' differently, as 'does not both express and possess a property' then it is heterological.)

Russell's Paradox involves the predicate (RP) 'class or set which is not a member of itself.' It seems that RP expresses a property, and since it is a truth of logic (call it the Abstraction Principle) that to every property there corresponds the class of all and only the things having that property, the Russell predicate, through expressing that property, determines a class which, it seems, can neither belong nor fail to belong to itself.

Cantor's paradox involves the same Abstraction Principle applied to the property of being a thing to yield (UC) the Universal Class. Cantor's Theorem says that every class is of lower cardinality than its power class (the class of all its subclasses) which implies that UC is of lower cardinality than its power class PUC. But this is incompatible with the requirement that any member of PUC must of course, like everything else, belong to UC.

RP is logically similar to 'heterological.' To say (truly) that the class of men does not belong to itself is to say it is not a man. To say (falsely) that the class of classes is not a member of itself is to say that it is not a class. There is no such property as being a non-self-membered class and thus no such class. This does not at all impugn the Abstraction Principle. There are indeed other ways of forming classes than as the extensions of properties, but they are inadequate for the determination of classes on the scale of interest to mathematical study. Mental acts of attention can identify a class, but not a very big one. Some will appeal to the mental powers of God, but they are especially unsuited to the task of forming classes by mental attention, since God is equally and perfectly aware of absolutely everything. God distinguishes things not by paying special attention but by knowing what properties they have.

Versions of Cantor's Theorem in first order set theories are unassailable specimens of mathematical truth. But as a principle of philosophy, it is false. Its proof depends on an alleged class essentially the same as the Russell Class. It is assumed that there is a one-to-one mapping M between UC and PUC. Then it is held there would have to be a class URC of all elements of UC which have the property RUP of not belonging to their M-correlate from PUC. URC would have to be a member of PUC and have an M-correlate X in UC. Now X is a member of URC if and only if it is not a member. The situation is exactly like Russell's Paradox.

The answer should also be the same. There is no such property as RUP for the same reason that there is no such property as RP. If there is such a property as existing (which has, of course, been disputed), then the Abstraction Principle guarantees the existence (and self-membership) of UC. Among its peculiarities will be its isomorphism with its power class. This is more satisfactory philosophically than making it out to be a thing which does not itself belong to any class. It would be better to say that Cantor's Theorem only holds for 'sets' and that UC is thus not a set in that sense.

One modal liar, (ML) is

(E) The proposition E expresses is not a necessary truth.

It seems to be (contingently) true that the proposition that E expresses is that the proposition that E expresses is not a necessary truth. It seems that proposition could not fail to be true. For if it were not true that that proposition is not necessary, it would be necessary, which is incompatible with its not being true. But then, since it cannot fail to be true, it must be necessary. But that implies that it is false.

Here it is best to answer that what E, as a matter of contingent fact, says, is that what it says is nonnecessary. Since what it says is (at least) that it is nonnecessary, it must then say that it is nonnecessary that it is nonnecessary. But this is, in the broad sense, a contradiction, since it is necessarily true that if anything is nonnecessary, then it is necessarily true that it is nonnecessary. That is the characteristic axiom of S5. Its utility in this case is just one more indication of its truth.

Yablo's infinite liar (YL), in one version, involves the sentence form

(F) Every sign along The Path numbered n or greater expresses a falsehood, where a sentence of that form is written on each of an infinite series of signs arranged in such a way that each one points in the same direction along The Path and is labeled with a number one less than the value of n in its sentence. (This has been held to have the consequence that not a one of the signs in this series is self-referential.) Now, if any one of the signs, X, were true, then every sign following it would be false. But then any sign Y following X would be such that every sign following Y was false, which would make Y true. Thus the truth of any sign X in the series entails a contradiction. So all the signs in the series would have to be false. But that seems to entail that each sign in the series would be true.

YL essentially involves the idea of a completed infinite series. If the signs were being produced one a day into the indefinite future, there would be no basis for trouble. A plain ordinary falsehood might turn up at any time, which would make the sentence immediately before it unproblematically true and thus in turn, all the sentences preceding it unproblematically false. This might seem implausible if the signs are being produced by a machine that merely puts up a duplicate sign a step down The Path each day, with nothing in the machine's repertoire to allow it to do anything else. But infinite time allows all sorts of things to happen. If it is added to the specifications that the machine is not going to break down, it will be a primary question whether this guarantee can be accommodated by a merely potential infinite. But if the complete infinite series could exist, the above argument for a paradoxical contradiction would apply.

This has been seen by some as showing that logical paradox of the Liar type does not depend on self-reference. As was observed above, this would depend on what is meant by self-reference. In any case, that question would not be important on the present approach, since this case of YL can be treated in the same way as EL. For any n, sign n attributes falsity to what is said by sign n + 1 and to what is said by sign n + 2. But n + 1 attributes falsity to what is said by n + 2. So n attributes falsity to what is said by n + 2 and also attributes falsity to that attribution of falsity. Thus sign n contradicts itself for each n. This is assuming the series is infinite. (If it is not, the result is different, but that need not be considered now.)

One Knower family paradox is (UK):

(G) No one knows that H is true.

This seems probably true by showing that the assumption to the contrary leads to a contradiction. And yet giving this proof somehow cannot qualify anyone as knowing H is true. Variations on this theme have been offered as *the* 'Surprise Test Paradox.' A teacher announces "There will be a test tomorrow and none of you know that this announcement is true." There are actually a number of candidates for paradox about 'surprise' events and some involve no announcement at all, just a known tradition of the 'teacher' being punished if he fails to spring a test which qualifies as a 'surprise,' and related arguments suggesting that he can (and cannot) succeed.

Paradoxes of the Knower family are not resolvable by the method used above for paradoxes of assertion and predication. Attributing having an unknown truth value is not like attributing truth or falsity or necessary truth or the like. These paradoxes need to be treated in a way similar to the belief case C above. They are cases requiring talking with alleged believers rather than cases which involve merely the logical powers of sentences by themselves. Does the assertor of H really believe that he does not know that what he is saying is true? Consider

(G′) No one believes that G′ is true.

A foreigner could easily believe that G′ expresses a true proposition, thanks to being ignorant as to what proposition it does express. But is there a proposition G′ expresses to intelligent speakers of English, which none of them believe? If not, should we not then conclude that no one believes, with full understanding, that G′ is true? And then, does not that proposition appear to be, after all, what one with a full understanding of G′ would see it as expressing? And doesn't that put us in logical trouble? The mistake here is in thinking that a proposition that turns up in the course of a certain line of reflection on G′ could have been the one G′ by itself was expressing all along. No one can believe G′ in a certain self-referential way. One can use G′ to express this thought. But then, so used, G′ is being believed true. To take that as proof that G′ is false is to slip into a confused equivocation. It is not the property of truth that needs stratifying here, but the various thoughts that get associated with G′.

A more serious problem about the applicability of the present approach can be brought out by considering a case in which a dozen people are required by the law to make exactly one deposition in regard to a certain case, and each one of them deposes a token of

(H) Something deposed by one of the others is false.

This case can be presented without reference to people, in terms of sentences by themselves, so that our rules about assertion and predication should be the answer. We could arbitrarily stipulate that a certain one of these depositions is false. That would have the consequence that all the others are true, which works out nicely as far as consistency goes. But this arbitrariness is obviously unacceptable. Here the above rule R4, which is the basis for answering the Grelling paradox, is not adequate, because no one of the depositions is such that no other of the depositions is a false one.

R4 is also not applicable to a version of Yablo's paradox in terms of existential quantification, in which the signs read

(I) Some sign along The Path, numbered n or greater, expresses a falsehood,

with each sign, as before, labeled with a number one less than the value of n in its sentence. This is unproblematic if the series is finite, since the last sentence is then false, making its predecessors true. But the infinite series raises the problem that if any one of the signs were false, all its successors would have to be true, which is impossible, leading to a contradiction just as with G. However, the treatment which works for YL-G and EL does not work for YL-I. And R4 does not work either, for the same reason that it does not work for H.

It might be tempting to write off YL-I as just a paradox of the completed infinite. Whether such paradoxes are logical paradoxes is the sort of question set aside at the beginning of our discussion. It would depend on whether it is a truth of logic that there are completed infinites. But we can continue to spare ourselves this question by noting the similarity between the problem of YL-I and that of the obviously finite case H. The inadequacy of R4 is the same for each.

Let us proceed directly to R5: To assert that some F is a G is to assert that if anything is such that nothing other than it is any better candidate for being an F that is a G than it is, then it is an F that is a G. R5 deals nicely with H. Every one of the deposers i has said of every other one of the depositions [{1,2, . . . 12} – i] that it is false, since every member of [{1,2, . . . ,12} – i] is equally qualified for being a deposition other than i which is false. But calling any one of these, j, false, is to attribute falsity to the claim that some one of [{1,2, . . . ,12} – j] is false which is to endorse all those claims, while calling all of them other than i false. Thus all the depositions are inconsistent and for that reason all are false. And that is not sufficient to make any of them true, because they have not claimed simply that one of them is false, as an external observer could simply claim.

The same goes for YL-I. For any sign n, all the subsequent signs qualify equally as candidates for a false subsequent sign and are thus, by R5, all called false by n. This makes each sign n contradictory just as in the case of YL-G. They do not claim simply that a subsequent is false, as an external observer could.

It will be objected that R5 does not sound at all like a logical rule, but more like something from ethics. The notion of being as good a candidate as there is, for being an F that is a G, will be held to be objectionably vague. It may well be vague in many cases. But in the two problem cases just considered it is perfectly clear. The problems in fact arose from the fact that it would be absurdly arbitrary to treat one candidate as a better case of an F that is a G than any among a set of others.