HILARY PUTNAM

THERE IS AT LEAST ONE A PRIORI TRUTH

In a number of famous publications (the most famous being the celebrated article 'Two Dogmas of Empiricism', *Philosophical Review*, 1951) Willard van Orman Quine has advanced the thesis that there is no such thing as an (absolutely) *a priori* truth. (Usually he speaks of 'analyticity' rather than apriority; but his discussion clearly includes both notions, and somewhere – I don't have the reference at the moment – he has explicitly said that what he is rejecting is the idea that any statement is completely *a priori*. For a discussion of the different threads in Quine's arguments, see my paper 'Two Dogmas Revisited', in *Contemporary Aspects of Philosophy*, ed. Gilbert Ryle, Oriel Press, 1977). Apriority is identified by Quine with *unrevisability*. But there are at least two possible interpretations of unrevisability: (Interpretation One) A *behavioral* interpretation, *viz.* an unrevisable statement is one we would never give up (as a sheer be-havioral fact about us); and (Interpretation Two) an *epistemic* interpreta-tion, *viz.* an unrevisable statement is one we would never be *rational* to give up (perhaps, even a statement that it would never be rational to even *think* of giving up). On the first interpretation, the claim that we might revise even the laws of logic becomes merely the claim that certain phenomena might cause us to give up our belief in some of the laws of logic; there would be no claim being made that doing so would be rational. Rather the notion of rationality itself would have gone by the board.

I don't know if Quine actually intended to take so radical a position as this, but, in any case, I think that most of his followers understood him to be advocating a more moderate doctrine. This more moderate doctrine was, in any case, put forward by me, for example, in a paper I titled 'It Ain't Necessarily So'. The moderate doctrine, unlike the more radical doctrine, employs the notion of rationality. The claim of the moderate doctrine is that there are no truths which it would never be rational to give up; for every truth or putative truth, there are circumstances under which it would be rational to accept its denial. This position was itself argued for, on the basis of an induction from the history of science. It was not itself supposed to be an *a priori* truth. Thus the cheap shot, which

consists in arguing that the anti-apriorist position is self-refuting because
if it were correct then there would still be one *a priori* truth, namely that
there are no *a priori* truths, doesn't work. But the induction from the
history of science was a somewhat complicated affair. It was not a simple
Baconian induction; rather, a theory was put forward, a theory which
was intended, among other things, to explain why certain statements *seem*
to be *a priori*.

I want to emphasize this point. The moderate Quinean position tries
to 'save the appearances'. It does not deny that there at least appear to
be *a priori* truths, it does not deny that certain truths have a special status,
it tries to explain why that is so. More precisely, it says that those truths
really do have a special status, only the status has been misconceived. The
key notion here was the notion I called 'contextual apriority'. The idea
is that we can grant that certain truths, and even, at certain times, certain
falsehoods, have a special status, but that we don't have to concede that
that status is good old-fashioned apriority. The status these truths and
falsehoods have, as long as they have it, is contextual apriority – apriority
relative to the body of knowledge. And the thesis that there are no *a priori*
truths becomes the thesis that there are no absolutely *a priori* truths.
What still seems to me to be right about this is the idea that there is such
a status as contextual apriority, and the idea that contextual apriority has
sometimes been mistaken for absolute apriority, that is, for the status that
a statement has if indeed it could never be rational to revise it.

There is an important difference between such statements as "the leaves
always turn in October", which can be refuted by just well-confirmed
observations, and such statements as the statements which comprise
non-Euclidean geometry as a theory of actual space (or space-time) which
can only be established when a whole new body of theory, not just geo-
metrical theory, but physical theory and experimental interpretation, is
put forward. Prior to the development of general relativity theory, most
people, even most scientists, could not imagine any experiences that would
lead them to give up, or that would make it rational to give up, Euclidean
geometry as a theory of actual space; and this is what led to the illusion
that Euclidean geometry was *a priori*. What I no longer think is that all
cases of apparent apriority can be explained in this fashion.

Even the case of Euclidean vs. non-Euclidean geometry involves features
that were glossed over in my previous account. It is not the case that every

mathematician regarded non-Euclidean geometry as *a priori* impossible, as a description of actual physical space, prior to the development of general relativity. Indeed, Lobachevskii always regarded the question of which geometry describes actual physical space as an empirical question. And it isn't just the possibility of giving an operational interpretation to non-Euclidean geometry that is important, although this was naturally stressed by empiricists like Lobachevskii but it is also important that one can give a coherent model for a non-Euclidean world within Euclidean mathematics. Mathematicians were led by a very straightforward analogy to grant the conceivability of Euclidean spaces of four and even more dimensions. A three-dimensional, non-Euclidean world – or at least a world whose intrinsic geometry, whose geometry viewed from within, is that of a three-dimensional non-Euclidean world, can be pictured as a curved hyper-surface in a four-dimensional Euclidean space. Of course, this doesn't explain the possibility of a non-Euclidean world which is *not* embedded in a higher-dimensional Euclidean space!

What I want to do today is to argue that there is at least one *a priori* truth in exactly the sense that Quine and I denied; that is, at least one truth that it would never be rational to give up. My example, not surprisingly, is going to be taken from the laws of logic. In the past I have argued that the laws of logic are revisable and that, in fact, the proper interpretation of quantum mechanics requires that we give up the distributive laws. Nothing that I say today will go against this position. It is after all perfectly possible that not all the traditional laws of logic are *a priori*, but that only some of them are. Indeed, even if, as I think, the notion of apriority has to be revived, that does not mean that we should go back to the old confident way of using it. To try to understand the epistemology of all of logic and classical mathematics in terms of a single notion of *a priori* truth would be, I think, a serious mistake. The law of logic I want to consider is a very weak version of the Principle of Contradiction. The Principle of Contradiction says that no statement is both true and false, or in the notation of propositional calculus, $\sim(p. \sim p)$.

The example of quantum logic suggests one way in which the revision of this principle might be suggested. Namely, it might be suggested that the principle holds only for *ordinary* statements about *ordinary* macro-observable properties of *ordinary* macro-observable objects, e.g., 'the cat is on the mat', and it might be suggested that there is some class of

recherché statements about waves and particles or whatnot for which the principle fails. Perhaps 'the electron is a particle' is both true and false, or 'the electron is a wave' is both true and false. This move might be avoided by considering what we may call the *Typical* Principle of Contradiction, that is, the principle that *ordinary macro-observable statements*, ordinary statements about macro-observables, *are not both true and false*, or by considering the principle that *most statements are not both true and false*, or some combination of these moves; but I shall consider the weakest possible version of the Principle of Contradiction, which I shall call the Minimal Principle of Contradiction. This is simply the principle that *not every statement is both true and false*. The denial of this principle is, of course, the claim that *every statement is both true and false*. If every statement is such that under some circumstances it might be rational to revise it, then under some circumstances it might be rational to accept that *every statement is both true and false*. Is this the case? Well, it certainly doesn't seem to be the case. And if it is not the case, if, indeed, there are no circumstances under which it would be rational to give up our belief that *not every statement is both true and false*, then there is at least one *a priori truth*. And one is all we need.

My argument is in this respect like Descartes'. I believe that one of the several things that Descartes wanted to do with his *cogito* was to establish precisely that there are *a priori* truths. And for the purpose of making this point, one needs only one example. Is, then, the statement that not every proposition is both true and false not an example of an absolutely, unconditionally, truly, actually *a priori* truth?

Recall that part of the strategy of what I called the moderate Quinean position was to save the appearances by showing that what we mistake for absolute apriority is a status which some propositions truly have, a status which is truly different from ordinary, garden-variety contingency, but which is not an absolute apriority. This is the status of contextual apriority. Is it possible that the Minimal Principle of Contradiction is then only a contextually *a priori* truth which we are tempted to mistake for an absolutely *a priori* truth?

The suggestion would be this: that there is some weird physical theory T which we have not yet thought of, but which implies the denial of the Minimal Principle of Contradiction and that someday when some scientist – some future Einstein – invents the theory T and shows us what

beautiful predictions it leads to, and how much it enhances our under-
standing and control of nature to accept the theory T, then we will all be
converted and by a kind of 'gestalt switch' we will go over to accepting
the theory T and to denying the Minimal Principle of Contradiction.

But there is an obvious problem with this line. The problem is that it's
quite obvious what the theory T will have to be. If we ever give up the
Minimal Principle of Contradiction, that is, if we ever come to believe
that every statement is both true and false, then its perfectly obvious what
the theory T will have to be. The theory T will have to be the theory which
consists of every statement and its negation! That is the theory T will
have to consist of such statements as 'the earth is round', 'the earth is not
round', 'two and two are four', 'two and two are not four', 'the moon is
made of green cheese', 'the moon is not made of green cheese', 'there are
quarks', 'there are no quarks', etc., etc., etc.... For once we are in the
happy position of being able to say exactly what the 'surrounding theory'
will have to be if we come to revise a particular contextually *a priori*
statement.

Of course, my move here might be challenged. One might, for example,
suggest that we will give up the Minimal Principle of Contradiction and
the Law of Double Negation at the same time. Then we might accept *it
is not the case that it is not the case that every statement is both true and
false*, without accepting that every statement is both true and false.
However, in that case the statement 'every statement is both true and
false' would still have the status of being *a priori* false, even if the state-
ment of which it was the negation isn't *a priori* true. And to concede the
existence of such a status as *a priori* falsity is, I think, as much as to
concede the existence of such a status as *a priori* truth. I assume, there-
fore, that I am dealing with an opponent who maintains not merely that
we might accept the double negation of the statement that *every statement
is both true and false*, but that we might accept that statement itself.

Again, it might be suggested that we will assert 'every statement is both
true and false', while at the same time giving up the Principle of Universal
Instantiation, which enables us to infer particular instances from an all-
statement. Then we would say the words 'every statement is both true and
false', but for no particular statement would we be committed to saying
of it that *it* is both true and false. But this would clearly be playing verbal
games. If I say the words 'every statement is both true and false', but I

don't conclude that 'the earth is round' is both true and false, or that 'two and two equals five' is both true and false, then I simply don't mean what is normally meant, or anything what is like what is normally meant by 'every statement is both true and false'.

In the case of geometry, when we went over to non-Euclidean geometry we didn't change the meaning of the words, or at any rate we didn't *merely* change the meaning of the words. We discovered that a state of affairs which we had mistakenly regarded as inconceivable is, in fact, conceivable and quite probably actual. For example, we used to regard it as inconceivable that a three dimensional world should be both finite and unbounded. We now think it is conceivable and quite probably the case that the whole three-dimensional universe is both finite and unbounded. The question is whether, in the same way, the state of affairs that we now regard as being inconceivable, the state of affairs that 'the earth is round' and at the same time 'the earth is not round', that 'the moon is made of green cheese', and at the same time that 'the moon is not made of green cheese', that 'two and two are five' and at the same time 'two and two are not five', and so on, is really conceivable and will perhaps someday turn out to obtain. Could it be rational to think someday that 'the moon is made of green cheese', and 'the moon is not made of green cheese', that 'two and two are five', and 'two and two are not five', that 'the earth is round', and 'the earth is not round', and so on? That is our question. And I repeat, if that ever happens then we know exactly what the 'theory' will be that we shall be accepting. It will have to consist of every statement and its negation.

Let me refer to the statement that *Euclidean geometry is true* – the statement we gave up when we went over to non-Euclidean physics – as the *critical statement*, and to the theory of the basis of which we decided that the critical statement was false – the General Theory of Relativity – as the *Embedding Theory*. What I've said so far is that if we take the Minimal Principle of Contradiction as our critical statement, then we know exactly what the Embedding Theory has to be. It has to consist of every statement and its negation. But it may still be argued that there is a disanalogy between accepting non-Euclidean geometry on the basis of the General Theory of Relativity and accepting the denial of the Minimal Principle of Contradiction on the basis of the theory which consists of every statement together with its negation. The disanalogy is that the

General Theory of Relativity leads to testable predictions, whereas the Embedding Theory which consists of every statement together with its negation leads to no testable predictions. But this is not the case either. The Embedding Theory in the latter case leads, for example, to the prediction that 'my hand has five fingers', and to the prediction that 'my hand has seven fingers'. It also leads to the prediction that 'my hand does not have five fingers', and to the prediction that 'my hand does not have seven fingers'. It leads to a *lot* of predictions! But, it may be objected, these are not genuine predictions for we don't know what it would be like for them all to come true. We can imagine all of the predictions of non-Euclidean physics coming true, even if we happen to be Euclidean physicists. But we don't know what it would be like for all the predictions of the theory that consists of every statement together with its negation to come true. I think this is right, but I think that this observation only poses the problem of apriority and does not solve it.

(Takes out box) In this box there is a sheet of paper. Suppose I predict that when I open the box you will see that the sheet of paper is red, and the sheet of paper is not red. Suppose I explain that I don't mean that the sheet of paper is red on one side and white on the other side, or anything like that. When I say that the sheet of paper is red, I mean that it's red on both sides – a nice, normal dye which doesn't look red from one angle and some other color from a different angle, or red to some people and a different color to other people, or anything like that. And when I say that the sheet of paper is red, and the sheet of paper is not red, I mean that the statement that 'the sheet of paper is red' understood as just indicated, is both definitely true and definitely false. Now it's quite true that in a certain sense we don't know what it would be like for that prediction to be verified, and that's our reason for denying that it is a genuine prediction about what will be seen when the box is opened. But one has to be careful here.

The kind of inconceivability that is relevant is not mere unintuitability. Let me say that we can intuit a state of affairs if we can actually visualize it. (I want to stick to a notion of intuition that's close to perception.) Now, we can predict that something will happen which we cannot intuit, al-although we can, in a sense, conceive of it happening. For example, I might predict that when I open the box you will see that the sheet of paper is a shade of red that none of you has ever seen. I think that you'd all

accept that as a perfectly good prediction, even though you can't intuit
what it would be like for that prediction to come true. It's enough that
we should be sure that that's a possible state of affairs or at least a state
of affairs that we could recognize if it turned out to be actual. Similarly,
if I predict that when I open the box you will see that the sheet of paper
is a color – and I mean now a major color – that you've never seen before,
I think that that would be a perfectly good prediction. It's true that such
a prediction would upset a certain amount of physical theory, namely the
physical theory that says that color is determined by lambda, the wave-
length of the light reflected from the paper. For if that theory is true,
and it's also true that we've correctly mapped out which lambdas corre-
spond to which colors, and which lambdas the human eye is sensitive to,
then there is no room, in the sense of no room in the theory, for another
major color. Nevertheless, it would be absurd to say that someone who
predicted that there was another major color and who claimed to have
predicted that when we opened the box and looked at the paper we would
see a major color we hadn't seen before, hadn't made a prediction just
because we couldn't intuit the state of affairs that would obtain if his
prediction turned out to be correct.

Actually the situation is more complicated than I'm suggesting because,
in fact, the physical theory that I just mentioned, although it still appears
in many textbooks, is certainly false, and the work of Jerome Letvin and
of Irwin Land shows that color depends in a very complicated way on
many factors besides lambda, and as far as I know it would not be the
case that the discovery of a new major color tomorrow would very much
mess up physical theory – there just isn't a good physical theory of color
to mess up. For example, standard theory doesn't really account for the
color "brown". But even if the lambda theory were not already suspect,
I think that the fact remains that the prediction of a new major color
would have to be counted as a prediction, even if we knew that verification
of that prediction would mess up a certain amount of well-established
theory.

Now, what do we mean when we say that we don't understand what it
would be like for the prediction, that when I open this box you will see
that the sheet of paper that it contains is both red and not red, to turn
out to be true? We mean at least that we cannot intuit what it would be
like for an observational situation to obtain which would *clearly* be

describable by saying that the sheet of paper is red, in the sense I explained before, and also the sheet of paper is not red; but we had better mean more than that, otherwise this counts as a perfectly good prediction. Just as the sheet of paper is a shade of red that you have never seen before, and the shade of paper is a major color that you have never seen before both count as perfectly good predictions.

On the other hand, it isn't that 'the sheet of paper is red' and 'the sheet of paper is not red' is literally unintelligible in the way in which 'wa'arobi besnork gavagai' is literally unintelligible, although some philosophers have tried to assimilate the unintelligibility of contradictions to the unintelligibility of what is literally without sense in the language. 'This sheet of paper is red and this sheet of paper is not red' isn't unintelligible at all. It simply asserts what cannot possibly be the case. And the reason that *when I open the box you will see that the sheet of paper is red and the sheet of paper is not red* does not count as a prediction, is that we know – know *a priori* – that it can't possibly turn out to be the case. But this remark doesn't explain the phenomenon of *a priori* knowledge, it only points to its existence.

If what I've said so far is correct, then the theory that what is happening, what gives rise to the illusion of apriority, is that we mistake one status for another – mistake the status of contextual apriority for the status of absolute apriority – doesn't work in this case. That was what was going on in the case of non-Euclidean geometry. But to explain the special status of the Principle of Contradiction, or at least of the Minimal Principle of Contradiction, in terms of contextual apriority, is a loser.

At this point there is a rather tough line that we might take. We might say that if every statement is both true and false, then in particular the statement 'my hand has five fingers' (or your favorite observation report) is both true and false. But I see that my hand has five fingers is true and I see that it is not false. So I *observe* that at least one statement is not both true and false, and this is enough to verify the Minimal Principle of Contradiction. This is a tough line to take because it amounts to giving up the search for a *special status* for the minimal Principle of Contradiction. It amounts to saying that the Minimal Principle of Contradiction is an *observation report* or is grounded upon a number of observation reports. But this is clearly wrong. It might turn out that there are not five fingers on my hand. For example, my hand may have been amputated and what

I'm looking at may be a plastic substitute (of course we'd have to tell some story about why I don't realize that I'm not looking at my own hand, but that is not impossible, as we all know). But even if it turned out that I don't have a hand, or that my hand has only four fingers, or seven fingers, or whatever, discovering that I was wrong about the observation report would not at all shake my faith in my belief that that observation report is not both true and false. Even if I couldn't discover how many fingers there are on my right hand (imagine a drunken man looking at his hand), this would not shake my faith in my belief that it's not both true and false that the number is five. We seem to be struck with at least one *a priori* truth – really, actually, truly *a priori*, and not just contextually *a priori*.

If we cannot successfully deny that there are *a priori* truths then it has seemed to many that we can give a conventionalist account of how *a priori* truth is possible. According to a typical such account, it is simply a *rule of language* that one must not assert both a statement and its negation, or to ascend to the meta-language, that one must not apply both the predicates "true" and "false" to the same statement. Moreover, these rules are seen as constituting the meanings of negation and of falsity, or as partially constituting the meanings of negation and of falsity, respectively. Anyone who both asserts a sentence and its syntactic negation other than for special purposes, e.g., to call attention to an ambiguity in the situation, is going against the meaning of the negation idiom. Thus if I say 'It is raining and it is not raining', and I don't mean simply to call attention to the fact that the particular situation leaves some room for discretion in the application of the description 'it is raining', or something of that kind, then I am going against the meaning of the words. And this is *why* the Principle of Contradiction is correct.

This account has a very fundamental defect which seems, strangely, not to have been noticed. *It explains much too much*. The problem with this account and with a number of other attempted accounts is that if it were correct, it wouldn't merely explain the status of the Principle of Contradiction in our *knowledge*, it would explain the Principle of Contradiction itself. It wouldn't just provide a reason that we know the Principle of Contradiction, it would provide a reason that the Principle of Contradiction is true. But it is easy to see that there cannot be such a reason. The Principle of Contradiction is prior to anything that might be offered

as an explanation for its truth. For example, suppose the Principle of Contradiction were not true. Suppose that even the Minimal Principle of Contradiction were not true. Then every statement would be both true and false. Then of course it would be true that the Principle of Contradiction is true by convention. But it would also be true that the Principle of Contradiction is not true by convention. It would be true that our laying down certain linguistic stipulations does not cause the Principle of Contradiction to be true. To put it bluntly, you can't make the Principle of Contradiction true by convention unless it's *already* true. This objection, the objection of explaining too much, also applies to other historic empiricist attempts, and even non-empiricist attempts, to explain the laws of logic. For example, that they are the laws of thought, or that they arise from relations of our ideas.

Of course one might try a moderate conventionalism. That is, one might try saying that the laws of logic, or at least the Principle of Contradiction, or at least the Typical Principle of Contradiction, or at least the Minimal Principle of Contradiction are just true, and one might agree that the truth can't sensibly be explained in terms of anything else, but one might hold that what is a matter of convention is not the truth of these laws but their necessity or the rationality of believing them. This, however, does not seem very appetizing. To take the latter proposal first, if it's simply true by convention that it's rational to believe the laws of logic and this convention is simply the convention constituting the conventional use of the tri-syllabic English word "rational", then what we have is the somewhat notorious ordinary language solution to Hume's problem, only now proposed as the solution to the problem of deduction. With respect to the former, that is the appeal to ordinary language as a solution to Hume's problem, Wesley Salmon once remarked that all this amounts to is the claim that if you use induction then you have the right to apply to yourself the noise 'rational', *and isn't that nice*. Professor Strawson replied to Salmon by observing that our propensity to make inductions need not be thought of as either arbitrary on the one hand, nor as conventional on the other; it may be *natural*. I take it that by "natural" Strawson meant something like 'innate'. Now, whatever the virtue may be of regarding our propensity to make inductions as simply an innate tendency that we have, it does seem as if in this respect deduction is different from induction. To say that our faith in the most fundamental

principles of deductive logic, our faith in the Principle of Contradiction itself, is simply an innate propensity and that it has no need of justification just because it is an innate propensity, is to obliterate totally the distinction between reason and blind faith. Of course, I'm not accusing either Peter Strawson or David Hume of making this move; they would both restrict their nativist account to induction, and not deduction. Nor can I accept the view that the necessity of the laws of logic, that is the fact that they hold in all possible worlds not only in the actual world, the fact that even if we accept the laws of logic as true in the actual world, we cannot go on and say 'but of course they might not have been true', or at least we cannot say 'it might have been that every statement was both true and false', is accounted for by convention.

It is true that there are accounts of logical truths, notably Quine's, according to which such a schema as $\sim(p. \sim p)$, if valid at all, is *ipso facto* necessary, that is to say there's no difference on Quine's account between saying that every instance of $\sim(p. \sim p)$ is true in the actual world, and saying that it is necessary that $\sim(p. \sim p)$; but this seems to me to be wrong. For one thing this assumes what we may call a Humean account of the modalities, that is it assumes that what is true in possible worlds is totally determined by what is true in the actual world plus our conventions. If this is right then there cannot be two possible worlds in which the *same* events take place, but which are such that *if* a certain experiment had been performed, which never was performed in either world, then different things would have happened in the two worlds. Now ask yourself this question: Can you imagine two worlds in neither of which the experiment is performed. The experiment just requires too much energy and the government won't let the physicist use so much energy in one experiment. Exactly the same events happen in both worlds but it is the case that *if* the experiment had been performed, *if* a certain particle had been submitted to much, much higher energies than were ever concentrated in a small space, the in *one* of the two worlds the particle would have split and in the other it would not have split? In other words, does the totality of facts about what events *actually* take place determine the truth value of all statements of the form 'it is *possible* that p'? To me, at least, it seems that the answer is 'no', and if the answer is 'no', then both Quinean accounts of logical necessity and Humean accounts of causality have to be wrong. But I don't want to discuss this here, I simply want to point

out that anyone who shares my modal-realist intuitions has to reject the claim that the *necessity* of the principles of logic is any *more* a matter of convention than their truth is. If any one is tempted to hold it, the form of moderate conventionalism that consists in saying that the laws of logic are *just* true in the actual world, but that *given* that they're true in the actual world it's a matter of our convention that they're true in all *possible* worlds seems to me quite untenable.

Incidently, the claim that physical possibility statements are translatable into statements about what actually happens seems to me in no better shape than the claim that statements about material objects are translatable into statements about sense data; and if physical possibility statements are not disguised statements about what actually happens then it is hard to see how logical possibility statements can be. There is however an account which goes part of the way towards explaining the special status of at least some of the laws of logic. A version of this account was, I believe, offered by Saul Kripke in a seminar at Princeton in which he criticized my published views on quantum logic; and the root idea of the account is to be found already in Aristotle's remarks about the laws of logic.

The idea is that the laws of logic are so central to our thinking that they define what a rational argument is. This may not show that we could never change our mind about the laws of logic, that is that no causal process could lead us to vocalize or believe different statements; but it does show that we could not be brought to change our minds *by a rational argument*. Let me spell this out a little. Typical rational arguments either have the form of chains of deduction of the familiar 'if *a*, then *b*' form, or they have the form of inferences to the best explanation. But the latter sort of inductive arguments of the form 'if *a* then *b*'; '*b*, so probably or plausibly *a*' also rely on properties of the connective 'if then', specifically upon *modus ponens*. Both in inductive reasoning and in deductive reasoning we make use of the fact that our language contains a connective which satisfies transitivity and *modus ponens*. This does not show that these two rules of inference are separately or jointly unrevisable; but it does show that if somebody rejected both of them then we would have no way of arguing with him. And indeed, Aristotle remarks that if anyone pretends to disbelieve one of the laws of logic and undertakes to argue with us, we can easily convince him that his own argument presupposes the very laws of logic that he is objecting to.

Neither Aristotle nor Kripke make the mistake, however, of offering this account as an account of why the laws of logic are true in the first place. All this account says is that part of their very special epistemic character is explained by what Quine would call their *centrality*. That is, they're presupposed by so much of the activity of argument itself that it is no wonder that we cannot envisage their being overthrown, or all of them being overthrown, by rational argument. But we should be clear about what the centrality argument does not show. It does not show that a putative law of logic, for instance the Principle of Contradiction, could not be overthrown by *direct observation*. Presumably I would give up the Principle of Contradiction if I ever had a sense datum which was both red and not red, for example. And the centrality argument sheds no light on how we know that this could never happen.

NOTE

This is a first draft of a paper I never finished. I no longer agree with the conclusion for a number of reasons, but I think the arguments are still of interest. One way I would begin to meet some of the arguments in this paper is by distinguishing two senses of 'revise'. A statement may be 'revised' by *negating* it – e.g., saying 'this is not white', where formerly we said 'this is white'; or it may be revised by *challenging a concept it contains*. My present position – February 18, 1977 – is that there are statements that cannot be revised in the *first* way (in this I think the fore-going paper is completely right), but that every statement is eligible for revision in the second way.

The question raised in the last paragraph – how do we know that a *direct observation* might not in the future contradict the principle of Contradiction – assumes that *what we now say* and *what is the case* are totally independent. (The stance I referred to as 'metaphysical realism' in my Presidential Address to the A.P.A.) Even if we grant that we may in the future *say* 'this sheet of paper is white and this sheet of paper is not white', we don't have to grant that we might be *right*. It may be that under our *present* conceptual scheme it is *mandatory* to find some explana-tion of that *future* utterance under which it is *not literally correct*. In Quine's terminology, it may be that homophonic translation (taking the future utterances at 'face value') is inadmissible in this case. When I

wrote the foregoing paper, I would have replied: "even if we refuse to
say now that the future sheet of paper might be both red and not red,
that doesn't *of itself* make it true that the future sheet of paper *won't*
be both red and not red. How do we know it doesn't just make *us*
stubborn?" This assumes that there is an intelligible distinction *within*
our conceptual system between what it is possible to conceive of within
that system and what is really (independently of all conceptual systems)
the case. This is just what I criticize in the address referred to.

On the other hand, I am not urging that we regard all logical and
mathematical truth as simply the product of our translation-practices
(let alone of 'convention'). I have long urged that there is an irreducible
factual element in logic and mathematics (e.g., the *consistency* of a set
of conventions is not itself a convention); which is not to deny that
there is also a conventional *component* to logic and mathematics. I think
it is right to say that, within our present conceptual scheme, the Minimal
Principle of Contradiction is *so* basic that it cannot significantly be 'ex-
plained' at all. But that doesn't make it an "absolutely *a priori* truth",
in the sense of an absolutely unrevisable truth. Mathematical Intuition-
ism, for example, represents one proposal for revising even the Minimal
Principle of Contradiction – not by saying that it is *false*, but by denying
the applicability of the classical concepts of truth and falsity at all. Of
course, then there would be a *new* 'Minimal Principle of Contradiction' –
e.g., 'no statement is both proved and disproved' (where 'proof' is taken
to be a concept which does *not* presuppose the classical notion of 'truth'
by the Intuitionists); but this is not the Minimal Principle of Contradic-
tion. Every statement is subject to revision; but not in every *way*.

NOTE TO SUPERSEDE (SUPPLEMENT?) THE PRECEDING NOTE

Added December 23, 1977

As I continue to think about these matters, it now seems to me that the
preceding note does not do justice to what was *right* in the original paper.
Rather than simply revise it, I have chosen to supplement the original
paper-plus-note-which-I-added-later with yet *another* note for a meta-
philosophical reason: it seems to me, and it has also been remarked by
another philosopher I respect, that we philosophers are frequently torn

in just the fashion that I am torn now between opposing considerations, but we very infrequently show it in *print*. What we do is let ourselves be torn in private until we finally 'plonk' for one alternative or the other; then the published paper only shows what we plonked for, and not the being torn. For once, the present paper-plus-potentially-infinite-series-of-notes *will* show the 'being torn'.

The preceding note tried to rescue what I called the 'moderate Quinean' position by taking the line that "every statement is revisable but not in every *way*". Specifically, a distinction was drawn between giving up a statement by accepting its negation, and giving up a statement by giving up concepts which occur in the statement (as somehow defective).

I don't think this works. Consider the statement I used in the original paper to show that there exists *at least one a priori truth*. This was the statement: 'Not *every* statement is *both* true and false'. In the previous Note, I said we might give this up by giving up the *classical* notions of truth and falsity – e.g., by going over to Intuitionist logic and metatheory. But surely if we did *that* we wouldn't view it as *giving up* the concepts of truth and falsity; rather we would view it as giving up an incorrect *analysis* of those notions.

Here it seems Quine has an easy rejoinder. He can say "See! It's just as I told you. You can't draw a non-arbitrary line between changing the meaning of the words and changing collateral beliefs. And for that very reason you can't tell if the original statement is still being expressed by the sentence 'Not every statement is both true and false'. Lacking any meaningful notion of synonymy, that is of statement identity, the question of whether some *statement* (not, *sentence*!) is immune from revision lacks all sense."

But, as I have argued in the papers cited at the beginning, Quine isn't just arguing against the notion of *synonymy*. (If he were, then if linguists were to come up with a well-motivated proposal for assigning sentences to synonymy classes, Quine's work would lose all interest.) Much of Quine's argument – specifically, his historical argument from the succession of past scientific revolutions – was *independent* of the question of whether there is a good criterion for sentence synonymy. Quine excited philosophers because he put forward a picture of epistemology in which there was no *room* for apriority (miscalled 'analyticity' by Quine *and* his positivist opponents). He excited philosophers by putting forward a view

of epistemology in which "no statement is immune from revision" – a very different claim from the claim that the question, "Is every statement immune from revision?" is *meaningless*. It is this view of epistemology that I am now criticizing.

Moreover, we can finesse the question of whether adopting Intuitionism would or would not be to change the meaning of 'true' and 'false'. If it is true, as I argued in the preceding Note, that we can't give up the critical statement except by changing the meaning of 'true' and 'false' (i.e., "giving up the concepts"), then the following hypothetical must be *absolutely* unrevisable:

If the classical notions of truth and falsity do not have to be given up, then not every statement is both true and false.

(In general, as Gareth Evans once remarked to me, to say that a statement is revisable, but only in a certain way, is to say that a certain conditional is *un*revisable.)

Again, look at the situation the following way: Consider the following Rule of Inference (call it 'the Absolutely Inconsistent Rule'): *from any and all premise-sets, including the null set of premises, to infer every p*. The argument of the previous paper was that, whatever might be said about everything being up for revision in the big spiderweb (or field of force, or whatever your favorite metaphor may be) of beliefs, at least one thing is sure: it can never be rational to accept the Absolutely Inconsistent Rule. And this seems right.

Does this mean that we have to go back to the idea of fixed unchanging canons of rationality, which Quine so persuasively attacked? I don't think it does. It seems right to me that we use our scientific method to devise a better scientific method at every stage. (Reichenbach, who stressed this idea in much of his writing, compared this to all use of tools. The first crude tools were fashioned with our hands; then we used crude tools to fashion more refined tools, and so on.) We started with a 'method' which evolution has 'hard wired in' to our brains, and we used that 'method' to discover (after how many thousand of years?) some principles of deduction and induction, which, after more thousands of years, have begun to be explicitly formalized, at least in part, and to be ever more mathematically sophisticated. And these principles will undoubtedly guide us in the search for still better principles (together with the method 'hard wired in' to our brains, which we still have to fall back on more than we

like to admit). But the fact that the canons of rationality are themselves evolving doesn't mean they don't exist (*pace* Feyerabend, *pace* Foucault!), nor does it mean that, in the course of the evolution, *anything whatsoever* (including acceptance of the Absolutely Inconsistent Rule) might occur. Evolution, in the domain of *instruments,* doesn't imply total, protean, lack of definite structure.

But, after all, just how important is it that Quine is wrong in his *total* rejection of the *a priori*? In one way it is not very important. We do not have a good *theory* of rationality, and are unlikely to have one in the forseeable future. Lacking the 'rigid designator' of rationality, the theoretical definition which tells us what rationality is in every possible world (as 'water is H_2O' tells us what water is in every possible world), it is virtually hopeless to show with any semblance of good argument that any specific statement is such that it would be irrational to ever give it up (apart from special examples, such as the one I constructed). Nor do we really need a proof that a statement is *a priori* in this sense (rationally unrevisable) very often. If a statement has the property that *we cannot now describe* any circumstances under which it would be rational to give it up, that will surely suffice for most purposes of philosophical argument. But, if it is always dangerous to take on the burden of trying to show that a statement is absolutely *a priori*, the foregoing reflections show that it is not just dangerous but actually wrong to make the quick leap from the fact that it is dangerous to claim that any statement is absolutely *a priori* to the absolute claim that there are no *a priori* truths.

Harvard University