THE JOURNAL OF PHILOSOPHY

THE PROBLEM OF COUNTERFACTUAL CONDITIONALS [1]

I. THE PROBLEM IN GENERAL

THE analysis of counterfactual conditionals is no fussy little grammatical exercise. Indeed, if we lack the means for interpreting counterfactual conditionals, we can hardly claim to have any adequate philosophy of science. A satisfactory definition of scientific law, a satisfactory theory of confirmation or of disposition terms (and this includes not only predicates ending in "ible" and "able" but almost every objective predicate, such as "is red"), would solve a large part of the problem of counterfactuals. Accordingly, the lack of a solution to this problem implies that we have no adequate treatment of any of these other topics. Conversely, a solution to the problem of counterfactuals would give us the answer to critical questions about law, confirmation, and the meaning of potentiality.

I am not at all contending that the problem of counterfactuals is logically or psychologically the first of these related problems. It makes little difference where we start if we can go ahead. If the study of counterfactuals has up to now failed this pragmatic test, the alternative approaches are little better off.

What, then, is the *problem* about counterfactual conditionals? Let us confine ourselves to those in which antecedent and consequent are inalterably false—as, for example, when I say of a piece of butter that was eaten yesterday, and that had never been heated,

If that piece of butter had been heated to 150° F., it would have melted.

Considered as truth-functional compounds, all counterfactuals are of course true, since their antecedents are false. Hence

If that piece of butter had been heated to 150° F., it would not have melted

would also hold. Obviously something different is intended, and the problem is to define the circumstances under which a given

[1] Slightly revised version of a paper read before the New York Philosophical Circle, May 11, 1946. My indebtedness in several matters to the work of C. I. Lewis and of C. H. Langford has seemed too obvious to call for detailed mention.

counterfactual holds while the opposing conditional with the contradictory consequent fails to hold. And this criterion of truth must be set up in the face of the fact that a counterfactual by its nature can never be subjected to any direct empirical test by realizing its antecedent.

In one sense the name "problem of counterfactuals" is misleading, because the problem is independent of the form in which a given statement happens to be expressed. The problem of counterfactuals is equally a problem of factual conditionals, for any counterfactual can be transposed into a conditional with a true antecedent and consequent; e.g.,

Since that butter did not melt, it wasn't heated to 150° F.

The possibility of such transformation is of no great importance except to clarify the nature of our problem. That "since" occurs in the contrapositive shows that what is in question is a certain kind of connection between the two component sentences; and the truth of this kind of statement—whether it is in the form of a counterfactual or factual conditional or some other form—depends not upon the truth or falsity of the components but upon whether the intended connection obtains. Recognizing the possibility of transformation serves mainly to focus attention on the central problem and to discourage speculation as to the nature of counterfacts. Although I shall begin my study by considering counterfactuals as such, it must be borne in mind that a general solution would explain the kind of connection involved irrespective of any assumption as to the truth or falsity of the components.

The effect of transposition upon another kind of conditional, which I call "semifactual," is worth noticing briefly. Should we assert

Even if the match had been scratched, it still would not have lighted,

we would uncompromisingly reject as an equally good expression of our meaning the contrapositive,

Even if the match lighted, it still wasn't scratched.

Our original intention was to affirm not that the non-lighting could be inferred from the scratching, but simply that the lighting could not be inferred from the scratching. Ordinarily a semifactual conditional has the force of denying what is affirmed by the opposite, fully counterfactual conditional. The sentence

Even had that match been scratched, it still wouldn't have lighted

is normally meant as the direct negation of

Had the match been scratched, it would have lighted.

That is to say, in practice full counterfactuals affirm, while semifactuals deny, that a certain connection obtains between antecedent and consequent.[2] Thus it is clear why a semifactual generally has not the same meaning as its contrapositive.

There are various special kinds of counterfactuals that present special problems. An example is the case of "counteridenticals," illustrated by the statements

If I were Julius Caesar, I wouldn't be alive in the twentieth century,

and

If Julius Caesar were I, he would be alive in the twentieth century.

Here, although the antecedent in the two cases is a statement of the same identity, we attach two different consequents which, on the very assumption of that identity, are incompatible. Another special class of counterfactuals is that of the "countercomparatives," with antecedents such as

If I had more money, . . .

The trouble with these is that when we try to translate the counterfactual into a statement about a relation between two tenseless, nonmodal sentences, we get as an antecedent something like

If "I have more money than I have" were true, . . .

although use of a self-contradictory antecedent was plainly not the original intent. Again there are the "counterlegals," conditionals with antecedents that either deny general laws directly, as in

If triangles were squares, . . .

or else make a supposition of particular fact that is not merely false but impossible, as in

If this cube of sugar were also spherical, . . .

All these kinds of counterfactuals offer interesting but not insurmountable special difficulties.[3] In order to concentrate upon the

[2] The practical import of a semifactual is thus different from its literal meaning. Literally a semifactual and the corresponding counterfactual are not contradictories but contraries, and both may be false (cf. footnote 8). The presence of the auxiliary terms "even" and "still," or either of them, is perhaps the idiomatic indication that a not quite literal meaning is intended.

[3] Of the special kinds of counterfactuals mentioned, I shall have something to say later about counteridenticals and counterlegals. As for countercomparatives, the following procedure is appropriate:—Given "If I had arrived one minute later, I would have missed the train," first expand this to "$(\exists t)$. t is a time. I arrived(d) at t. If I had arrived one minute later than t, I would

major problems concerning counterfactuals in general, I shall usually choose my examples in such a way as to avoid these more special complications.

As I see it, there are two major problems, though they are not independent and may even be regarded as aspects of a single problem. A counterfactual is true if a certain connection obtains between the antecedent and the consequent. But as is obvious from examples already given, the consequent seldom follows from the antecedent by logic alone. (1) In the first place, the assertion that a connection holds is made on the presumption that certain circumstances not stated in the antecedent obtain. When we say

If that match had been scratched, it would have lighted,

we mean that conditions are such—i.e., the match is well made, is dry enough, oxygen enough is present, etc.—that "That match lights" can be inferred from "That match is scratched." Thus the connection we affirm may be regarded as joining the consequent with the conjunction of the antecedent and other statements that truly describe relevant conditions. Notice especially that our assertion of the counterfactual is *not* conditioned upon these circumstances obtaining. We do not assert that the counterfactual is true *if* the circumstances obtain; rather, in asserting the counterfactual we commit ourselves to the actual truth of the statements describing the requisite relevant conditions. The first major problem is to define relevant conditions; to specify what sentences are meant to be taken in conjunction with an antecedent as a basis for inferring the consequent. (2) But even after the particular relevant conditions are specified, the connection obtaining will not ordinarily be a logical one. The principle that permits inference of

That match lights

from

That match is scratched. That match is dry enough. Enough oxygen is present. Etc.

is not a law of logic but what we call a natural or physical or causal law. The second major problem concerns the definition of such laws.

have missed the train." The counterfactual conditional constituting the final clause of this conjunction can then be treated, within the quantified whole, in the usual way. Translation into "If 'I arrive one minute later than t' were true, then 'I miss the train' would have been true" does not give us a self-contradictory component.

II. The Problem of Relevant Conditions

It might seem natural to propose that the consequent follows by law from the antecedent and a description of the actual state-of-affairs of the world, that we need hardly define relevant conditions because it will do no harm to include irrelevant ones. But if we say that the consequent follows by law from the antecedent and *all* true statements, we encounter an immediate difficulty:—among true sentences is the negate of the antecedent, so that from the antecedent and all true sentences everything follows. Certainly this gives us no way of distinguishing true from false counterfactuals.

We are plainly no better off if we say that the consequent must follow from *some* set of true statements conjoined with the antecedent; for given any counterfactual antecedent A, there will always be a set S—namely, the set consisting of $-A$—such that from $A \cdot S$ any consequent follows. (Hereafter I shall regularly use "A" for the antecedent, "C" for the consequent, and "S" for the set of statements of the relevant conditions.)

Perhaps then we must exclude statements logically incompatible with the antecedent. But this is insufficient; for a parallel difficulty arises with respect to true statements which are not logically but are otherwise incompatible with the antecedent. For example, take

If that radiator had frozen, it would have broken.

Among true sentences may well be (S)

That radiator never reached a temperature below 33° F.

Now it is certainly generally true that

All radiators that freeze but never reach below 33° F. break,

and also that

All radiators that freeze but never reach below 33° F. fail to break;

for there are no such radiators. Thus from the antecedent of the counterfactual and the given S, we can infer any consequent.

The natural proposal to remedy this difficulty is to rule that counterfactuals can not depend upon empty laws; that the connection can be established only by a principle of the form "All x's are y's" when there are some x's. But this is ineffectual. For if empty principles are excluded, the following non-empty principles may be used in the case given with the same result:

Everything that is either a radiator that freezes but does not reach below 33° F., or that is a soap bubble, breaks;
Everything that is either a radiator that freezes but does not reach below 33° F., or is powder, does not break.

By these principles we can infer any consequent from the A and S in question.

The only course left open to us seems to be to define relevant conditions as the set of all true statements each of which is both logically and non-logically compatible with A where non-logical incompatibility means violation of a non-logical law.[4] But another difficulty immediately appears. In a counterfactual beginning

If Jones were in Carolina, . . .

the antecedent is entirely compatible with

Jones is not in South Carolina

and with

Jones is not in North Carolina

and with

North Carolina plus South Carolina is identical with Carolina;

but all these taken together with the antecedent make a set that is self-incompatible, so that again any consequent would be forthcoming.

Clearly it will not help to require only that for *some* set S of true sentences, $A \cdot S$ be self-compatible and lead by law to the consequent; for this would make a true counterfactual of

If Jones were in Carolina, he would be in South Carolina,

and also of

If Jones were in Carolina, he would be in North Carolina,

which can not both be true.

It seems that we must elaborate our criterion still further, to characterize a counterfactual as true if and only if there is some set S of true statements such that $A \cdot S$ is self-compatible and leads by law to the consequent, while there is no such set S' such that $A \cdot S'$ is self-compatible and leads by law to the negate of the consequent.[5] Unfortunately even this is not enough. For among true sentences will be the negate of the consequent: $-C$. Is $-C$ compatible with A or not? If not, then A alone without any additional condi-

[4] This of course raises very serious questions, which I shall come to presently, about the nature of non-logical law.

[5] Note that the requirement that $A \cdot S$ be self-compatible can be fulfilled only if the antecedent is self-compatible; hence the conditionals I have called "counterlegals" will all be false. This is convenient for our present purpose of investigating counterfactuals that are not counterlegals. If it later appears desirable to regard all or some counterlegals as true, special provisions may be introduced.

tions must lead by law to C. But if $-C$ is compatible with A (as in most cases), then if we take $-C$ as our S, the conjunction $A \cdot S$ will give us $-C$. Thus the criterion we have set up will seldom be satisfied; for since $-C$ will normally be compatible with A—as the need for introducing the relevant conditions testifies—there will normally be an A (namely, $-C$) such that $A \cdot S$ is self-compatible and leads by law to $-C$.

Part of our trouble lies in taking too narrow a view of our problem. We have been trying to lay down conditions under which an A that is known to be false leads to a C that is known to be false; but it is equally important to make sure that our criterion does not establish a similar connection between our A and the (true) negate of C. Because our S together with A was to be so chosen as to give us C, it seemed gratuitous to specify that S must be compatible with C; and because $-C$ is true by supposition, S would necessarily be compatible with it. But we are testing whether our criterion not only admits the true counterfactual we are concerned with but also excludes the opposing conditional. Accordingly, our criterion must be modified by specifying that S be compatible with both C and $-C$.[6] In other words, S by itself must not decide between C and $-C$, but S together with A must lead to C but not to $-C$. We need not know whether C is true or false.

Our rule thus reads that a counterfactual is true if and only if there is some set S of true sentences such that S is compatible with C and with $-C$, and such that $A \cdot S$ is self-compatible and leads by law to C; while there is no set S' compatible with C and with $-C$, and such that $A \cdot S'$ is self-compatible and leads by law to $-C$. As thus stated, the rule involves a certain redundancy; but simplification is not in point here, for the criterion is still inadequate.

The requirement that $A \cdot S$ be self-compatible is not strong enough; for S might comprise true sentences that although *compatible with A,* were such that *they would not be true if A were true.* For this reason, many statements that we would regard as definitely false would be true according to the stated criterion. As an example, consider the familiar case where for a given match M, we would affirm

(I) If match M had been scratched, it would have lighted,

[6] It is natural to inquire whether for similar reasons we should stipulate that S must be compatible with both A and $-A$, but this is unnecessary. For if S is incompatible with $-A$, then A follows from S; therefore if S is compatible with both C and $-C$, then $A \cdot S$ can not lead by law to one but not the other. Hence no sentence incompatible with $-A$ can satisfy the other requirements for a suitable S.

but deny

(II) If match M had been scratched, it would not have been dry.[7]

According to our tentative criterion, statement II would be quite as true as statement I. For in the case of II, we may take as an element in our S the true sentence

Match M did not light,

which is presumably compatible with A (otherwise nothing would be required along with A to reach the opposite as the consequent of the true counterfactual statement, I). As our total $A \cdot S$ we may have

Match M is scratched. It does not light. It is well made. Oxygen enough is present . . . etc.;

and from this, by means of a legitimate general law, we can infer

It was not dry

and there would seem to be no suitable set of sentences S' such that $A \cdot S'$ leads by law to the negate of this consequent. Hence the unwanted counterfactual is established in accord with our rule. The trouble is caused by including in our S a true statement which though compatible with A would not be true if A were. Accordingly we must exclude such statements from the set of relevant conditions; S, in addition to satisfying the other requirements already laid down, must be not merely compatible with A but "jointly tenable" or "cotenable" with A. A is cotenable with S, and the conjunction $A \cdot S$ self-cotenable, if it is not the case that S would not be true if A were.[8]

Parenthetically it may be noted that the relative fixity of conditions is often unclear, so that the speaker or writer has to make explicit additional provisos or give subtle verbal clues as to his meaning. For example, each of the following two counterfactuals would normally be accepted:

[7] Of course, some sentences similar to II, referring to other matches under special conditions, may be true; but the objection to the proposed criterion is that it would commit us to many such statements that are patently false. I am indebted to Morton G. White for a suggestion concerning the exposition of this point.

[8] The double negative can not be eliminated here; for ". . . if S would be true if A were" actually constitutes a stronger requirement. As we noted earlier (footnote 2), if two conditionals having the same counterfactual antecedent are such that the consequent of one is the negate of the consequent of the other, the conditionals are contraries and both may be false. This will be the case, for example, if every otherwise suitable set of relevant conditions that in conjunction with the antecedent leads by law either to a given consequent or its negate leads also to the other.

If New York City were in Georgia, then New York City would be in the South. If Georgia included New York City, then Georgia would not be entirely in the South.

Yet the antecedents are logically indistinguishable. What happens is that the direction of expression becomes important, because in the former case the meaning is

If New York City were in Georgia, and the boundaries of Georgia remained unchanged, then . . .

while in the latter case the meaning is

If Georgia included New York City, and the boundaries of New York City remained unchanged, then . . .

Without some such cue to the meaning as is covertly given by the word-order, we should be quite uncertain which of the two consequents in question could be truly attached. The same kind of explanation accounts for the paradoxical pairs of counteridenticals mentioned earlier.

Returning now to the proposed rule, I shall neither offer further corrections of detail nor discuss whether the requirement that S be cotenable with A makes superfluous some other provisions of the criterion; for such matters become rather unimportant beside the really serious difficulty that now confronts us. In order to determine the truth of a given counterfactual it seems that we have to determine, among other things, whether there is a suitable S that is cotenable with A and meets certain further requirements. But in order to determine whether or not a given S is cotenable with A, we have to determine whether or not the counterfactual "If A were true, then S would not be true" is itself true. But this means determining whether or not there is a suitable S_1, cotenable with A, that leads to $-S$ and so on. Thus we find ourselves involved in an infinite regressus or a circle; for cotenability is defined in terms of counterfactuals, yet the meaning of counterfactuals is defined in terms of cotenability. In other words to establish any counterfactual, it seems that we first have to determine the truth of another. If so, we can never explain a counterfactual except in terms of others, so that the problem of counterfactuals must remain unsolved.

Though unwilling to accept this conclusion, I do not at present see any way of meeting the difficulty. One naturally thinks of revising the whole treatment of counterfactuals in such a way as to admit first those that depend on no conditions other than the antecedent, and then use these counterfactuals as the criteria for the cotenability of relevant conditions with antecedents of other counterfactuals, and so on. But this idea seems initially rather unpromising in view of the formidable difficulties of accounting by such a step-by-step method for even so simple a counterfactual as:

If the match had been scratched, it would have lighted.

III. The Problem of Law

Even more serious is the second of the problems mentioned earlier: the nature of the general statements that enable us to infer the consequent upon the basis of the antecedent and the statement of relevant conditions. The distinction between these connecting principles and relevant conditions is imprecise and arbitrary; the "connecting principles" might be conjoined to the condition-statements, and the relation of the antecedent-conjunction $(A \cdot S)$ to the consequent thus made a matter of logic. But the same problems would arise as to the kind of principle that is capable of supporting a counterfactual; and it is convenient to consider the connecting principles separately.

In order to infer the consequent of a counterfactual from the antecedent A and a suitable statement of relevant conditions S, we make use of a general statement, namely, the generalization [9] of the conditional having $A \cdot S$ for antecedent and C for consequent. For example, in the case of

If the match had been scratched, it would have lighted

the connecting principle is

Every match that is scratched, well made, dry enough, in enough oxygen, etc., lights.

But notice that *not* every counterfactual is actually supported by the principle thus arrived at, *even* if that principle is *true*. Suppose, for example, that all I had in my right pocket on V–E day was a group of silver coins. Now we would not under normal circumstances affirm of a given penny P

If P had been in my pocket on V–E day, P would have been silver,[10]

[9] The sense of "generalization" intended here is that explained by C. G. Hempel in "A Purely Syntactical Definition of Confirmation," *Journal of Symbolic Logic*, Vol. 8 (1943), pp. 122–143.

[10] The antecedent in this example is intended to mean "If P, while remaining distinct from the things that were in fact in my pocket on V–E day, had also been in my pocket then," and *not* the quite different, counteridentical "If P had been identical with one of the things that were in my pocket on V–E day." While the antecedents of most counterfactuals (as, again, our familiar one about the match) are—literally speaking—open to both sorts of interpretation, ordinary usage normally calls for some explicit indication when the counteridentical meaning is intended.

even though from

P was in my pocket on V–E day

we can infer the consequent by means of the general statement

Everything in my pocket on V–E day was silver.

On the contrary, we would assert that if P had been in my pocket, then this general statement would not be true. The general statement will *not* permit us to infer the given consequent from the counterfactual assumption that P was in my pocket, because the general statement will not itself withstand that counterfactual assumption. Though the supposed connecting principle is indeed general, true, and perhaps even fully confirmed by observation of all cases, it is incapable of supporting a counterfactual because it remains a description of accidental fact, not a law. The truth of a counterfactual conditional thus seems to depend on whether the general sentence required for the inference is a law or not. If so, our problem is to distinguish accurately between causal laws and casual facts.[11]

The problem illustrated by the example of the coins is closely related to that which led us earlier to require the cotenability of the antecedent and the relevant conditions, in order to avoid resting a counterfactual on any statement that would not be true if the antecedent were true. But decision as to the cotenability of two sentences must depend upon decisions as to whether or not certain general statements are laws, and we are now concerned directly with the latter problem. Is there some way of distinguishing laws from non-laws among true universal statements of the kind in question, such that a law will be the sort of principle that will support a counterfactual conditional while a non-law will not?

Any attempt to draw the distinction by reference to a notion of causative force can be dismissed at once as unscientific. And it is clear that no purely syntactical criterion can be adequate, for even the most special descriptions of particular facts can be cast in a form having any desired degree of syntactical universality. "Book *B* is small" becomes "Everything that is *Q* is small" if "*Q*" stands for some predicate that applies uniquely to *B*. What then does distinguish a law like

All butter melts at 150° F.

[11] The importance of distinguishing laws from non-laws is too often overlooked. If a clear distinction can be defined, it may serve not only the purposes explained in the present paper but also many of those for which the increasingly dubious distinction between analytic and synthetic statements is ordinarily supposed to be needed.

from a true and general non-law like

All the coins in my pocket are silver?

Primarily, I would like to suggest, the fact that the first is accepted as true while many cases of it remain to be determined, the further, unexamined cases being predicted to conform with it. The second sentence, on the contrary, is accepted as a description of contingent fact *after* the determination of all cases, no prediction of any of its instances being based upon it. This proposal raises innumerable problems, some of which I shall consider presently; but the idea behind it is just that the principle we use to decide counterfactual cases is a principle we are willing to commit ourselves to in deciding unrealized cases that are still subject to direct observation.

As a first approximation then, we might say that a law is a true sentence used for making predictions. That laws are used predictively is of course a simple truism, and I am not proposing it as a novelty. I want only to emphasize the idea that rather than a sentence being used for prediction because it is a law, it is called a law because it is used for prediction; and that rather than the law being used for prediction because it describes a causal connection, the meaning of the causal connection is to be interpreted in terms of predictively used laws.

By the determination of all instances, I mean simply the examination or testing by other means of all things that satisfy the antecedent, to decide whether all satisfy the consequent also. There are difficult questions about the meaning of "instance," many of which Professor Hempel has investigated. Most of these are avoided in our present study by the fact that we are concerned with a very narrow class of sentences: those arrived at by generalizing conditionals of a certain kind. Remaining problems about the meaning of "instance" I shall have to ignore here. As for "determination," I do not mean final discovery of truth, but only enough examination to reach a decision as to whether a given statement or its negate is to be admitted as evidence for the hypothesis in question.

The limited scope of our present problem makes it unimportant that our criterion, if applied generally to all statements, would classify as laws many statements—e.g., true singular predictions—that we would not normally call laws.

A more pertinent point is the application of the proposed criterion to vacuous generalities. As the criterion stands, no conditional with an empty antecedent-class will be a law, for all its instances will have been determined prior to its acceptance. Now since the antecedents of the statements we are concerned with will be generalizations from self-cotenable and therefore self-compatible conjunctions, none will be known to be vacuous.[12] For example, since

M is scratched. *M* is dry . . . (etc.)

is a self-compatible set, the antecedent of

For every *x*, if *x* is scratched and *x* is dry (etc.), then *x* lights

will not be known to be false. But now we would still want the generalized principle just given to be a law if it should just *happen* to be the case that nothing satisfies the antecedent. This discloses a defect in our criterion, which should be amended to read as follows: A true statement of the kind in question is a law if we accept it before we *know* that the instances we have determined are *all* the instances.

For convenience, I shall use the term "lawlike" for sentences which, whether they are true or not, satisfy the other requirements in the definition of law. A law is thus a sentence that is both lawlike and true, but a sentence may be true without being lawlike, as I have illustrated, or lawlike without being true, as we are always learning to our dismay.

Now the property of lawlikeness as so far defined is not only rather an accidental and subjective one but an ephemeral one that sentences may acquire and lose. As an example of the undesirable consequences of this impermanence, a true sentence that had been used predictively would cease to be a law when it became fully tested—i.e., when none of its instances remained undetermined. The definition, then, must be restated in some such way as this: A general statement is lawlike if and only if it is acceptable prior to the determination of all its instances. This is immediately objectionable because "acceptable" itself is plainly a dispositional term; but I propose to use it only tentatively, with the idea of eliminating it eventually by means of a non-dispositional definition. Before trying to accomplish that, however, we must face another difficulty in our tentative criterion of lawlikeness.

Suppose that the appropriate generalization fails to support a given counterfactual because that generalization, while true, is unlawlike, as is

Everything in my pocket is silver.

[12] Had it been sufficient in the preceding section to require only that $A \cdot S$ be self-*compatible*, this requirement might now be eliminated in favor of the stipulation that the generalization of the conditional having $A \cdot S$ as antecedent and C as consequent should be non-vacuous; but this stipulation would not guarantee the self-*cotenability* of $A \cdot S$.

All we would need do to get a law would be to broaden the antecedent strategically. Consider, for example, the sentence

Everything that is in my pocket or is a dime is silver.

Since we have not examined all dimes, this is a predictive statement and—since presumably true—would be a law. Now if we consider our original counterfactual and choose our S so that $A \cdot S$ is

P is in my pocket. *P* is in my pocket or is a dime,

then the pseudo-law just constructed can be used to infer from this the sentence "*P* is silver." Thus the untrue counterfactual is established, if one prefers to avoid an alternation as a condition-statement; the same result can be obtained by using a new predicate such as "dimo" to mean "is in my pocket or is a dime." [13]

The change called for, I think, will make the definition of lawlikeness read as follows: A sentence is lawlike if its acceptance does not depend upon the determination of any given instance.[14] Naturally this does not mean that acceptance is to be independent of all determination of instances, but only that there is no particular instance on the determination of which acceptance depends. This criterion excludes from the class of laws a statement like

That book is black and oranges are spherical

on the ground that acceptance requires knowing whether the book is black; it excludes

Everything that is in my pocket or is a dime is silver

on the ground that acceptance demands examination of all things in my pocket. Moreover, it excludes a statement like

All the marbles in this bag except Number 19 are red, and Number 19 is black

on the ground that acceptance would depend on examination of or knowledge gained otherwise concerning marble Number 19. In fact the principle involved in the proposed criterion is a rather powerful one and seems to exclude most of the troublesome cases.

We must still, however, replace the notion of the acceptability

[13] Apart from the special class of connecting principles we are concerned with, note that under the stated criterion of lawlikeness, any statement could be expanded into a lawlike one; for example: given "This book is black" we could use the predictive sentence "This book is black and all oranges are spherical" to argue that the blackness of the book is the consequence of a law.

[14] So stated, the definition counts vacuous principles as laws. If we read instead "given class of instances," vacuous principles will be non-laws since their acceptance depends upon examination of the null class of instances. For my present purposes the one formulation is as good as the other.

of a sentence, or of its acceptance *depending* or *not depending* on some given knowledge, by a positive definition of such dependence. It is clear that to say that the acceptance of a given statement depends upon a certain kind and amount of evidence is to say that given such evidence, acceptance of the statement is in accord with certain general standards for the acceptance of statements that are not fully tested. So one turns naturally to theories of induction and confirmation to learn the distinguishing factors or circumstances that determine whether or not a sentence is acceptable without complete evidence. But publications on confirmation not only have failed to make clear the distinction between confirmable and non-confirmable statements, but show little recognition that such a problem exists.[15] Yet obviously in the case of some sentences like

Everything in my pocket is silver

or

No twentieth-century president of the United States will be between 6 feet 1 inch and 6 feet 1½ inches tall,

not even the testing with positive results of all but a single instance is likely to lead us to accept the sentence and predict that the one remaining instance will conform to it; while for other sentences such as

All dimes are silver

or

All butter melts at 150° F.

or

All flowers of plants descended from this seed will be yellow

positive determination of even a few instances may lead us to accept the sentence with confidence and make predictions in accordance with it.

There is some hope that cases like these can be dealt with by a sufficiently careful and intricate elaboration of current confirmation theories; but inattention to the problem of distinguishing between confirmable and non-confirmable sentences has left most confirmation theories open to more damaging counterexamples of an elementary kind.

Suppose we designate the 26 marbles in a sack by the letters of

[15] The points discussed in this and the following paragraph have been dealt with a little more fully in my "Query on Confirmation," this JOURNAL, Vol. XLIII (1946), pp. 383–385.

the alphabet, using these merely as proper names having no ordinal significance. Suppose further that we are told that all the marbles except *d* are red, but we are not told what color *d* is. By the usual kind of confirmation theory this gives strong confirmation for the statement

Ra. Rb. Rc. Rd. . . . Rz

because 25 of the 26 cases are known to be favorable while none is known to be unfavorable. But unfortunately the same argument would show that the very same evidence would equally confirm

Ra. Rb. Rc. Re. . . . Rz.—Rd,

for again we have 25 favorable and no unfavorable cases. Thus "*Rd*" and "– *Rd*" are equally and strongly confirmed by the same evidence. If I am required to use a single predicate instead of both "*R*" and "– *R*" in the second case, I will use "*P*" to mean:

is in the sack and either is not *d* and is red, or is *d* and is not red.

Then the evidence will be 25 positive cases for

All the marbles are *P*

from which it follows that *d* is *P*, which implies that *d* is not red. The problem of what statements are confirmable merely becomes the equivalent problem of what predicates are projectible from known to unknown cases.

So far, I have discovered no way of meeting these difficulties. Yet as we have seen, some solution is urgently wanted for our present purpose; for only where willingness to accept a statement involves predictions of instances that may be tested does acceptance endow that statement with the authority to govern counterfactual cases, which can not be directly tested.

In conclusion, then, some problems about counterfactuals depend upon the definition of cotenability, which in turn seems to depend upon the prior solution of those problems. Other problems require an adequate definition of law. The tentative criterion of law here proposed is reasonably satisfactory in excluding unwanted kinds of statements, and in effect, reduces one aspect of our problem to the question how to define the circumstances under which a statement is acceptable independently of the determination of any given instance. But this question I do not know how to answer.

NELSON GOODMAN

UNIVERSITY OF PENNSYLVANIA