Metaphysics texts: Rashi
Revised
18 Nov 2026
18:18
K S B

Fact, Fiction, *and* Forecast

Fourth Edition

NELSON GOODMAN

Harvard University Press
Cambridge, Massachusetts
and London, England

I

THE PROBLEM OF COUNTERFACTUAL CONDITIONALS[1]

1. *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. Conversely, a solution to the problem of counterfactuals would give us the answer to critical questions about law, confirmation, and the meaning of potentiality.

The thesis: the consequent follows by law from the antecedent plus relevant conditions. The conditions end in a circle (p. 16), law in an undefined 'acceptable' (p. 27).

disposition terms: *predicates for what a thing would do in given conditions. 'Soluble' means: would dissolve if put in water. Even 'red' says how a thing would look in daylight.*

Conversely: *the dependence runs both ways, so none of these notions can be taken as given in explaining the others. §3 bears it out: law leads back to confirmation, and stops there (p. 27).*

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.

[1] My indebtedness in several matters to the work of C. I. Lewis has seemed too obvious to call for detailed mention.

What pp. 4–5 settle: the truth-values of the clauses can't decide a counterfactual, so its truth must lie in a connection between them. Pp. 7–9 analyse that connection.

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.

truth-functional compounds: *read as the material conditional, which is false only when the antecedent is true and the consequent false. So a false antecedent makes it true whatever the consequent says.*

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 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.

opposing conditional: *same antecedent, negated consequent. This sets the test for every proposal in §2: it must pass the one and fail the other. Most proposals there are dropped for passing both.*

A point of method. The contrapositive has true clauses and poses the same problem. So what needs analysis isn't the false antecedent but the connection that is claimed.

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.,

transposed: *contraposed: 'if A then C' is turned into 'if not-C then not-A'. Since A and C are false, not-C and not-A are true, and the result is a factual conditional.*

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 statements of this kind—whether they have the form of counterfactual or factual conditionals 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.

a certain kind of connection: *'since' says one thing holds because of the other. So truth depends on a tie between the clauses, not on their truth-values. §§2-3 try to say what the tie is.*

counterfacts: *unrealized states of affairs for a counterfactual to describe. None are needed: the same problem arises for the contrapositive, which is about what did happen.*

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

Semifactuals (false antecedent, true consequent) confirm the diagnosis: 'even if' denies the connection a counterfactual affirms, so contraposing one changes its sense.

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.

Even if the match lighted: *the contrapositive, formed as on p. 4. It makes 'not scratched' follow from 'lighted', which nobody meant. So the original asserted no connection at all.*

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.

full counterfactuals affirm: *so both kinds of conditional are about one thing, whether the connection holds. An account of the connection, which §§2-3 look for, would then cover both.*

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.

Three special kinds, put aside so that §§2-3 can work on plain cases. Counteridenticals return on p. 16, counterlegals in note 5 on p. 12.

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.

the same identity: *'I am Caesar' and 'Caesar is I' say the same, yet the consequents clash. So more than the antecedent's content fixes what follows. On p. 16: the word order signals what is held fixed.*

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

Literally: *'even if A, still not C' then says that not-C follows from A. That and 'if A, C' can't both be true, but both are false when A, with the conditions, favours neither. 'Note I.9' is note 9, p. 15.*

[2] The practical import of a semifactual is thus different from its literal import. Literally a semifactual and the corresponding counterfactual are not contradictories but contraries, and both may be false (cf. Note I.9 below). 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.

two tenseless, non-modal sentences, we get as an antecedent something like

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

which wrongly represents the original antecedent as self-contradictory. 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, ...

Counterfactuals of all these kinds offer interesting but not insurmountable special difficulties.[3] In order to concentrate upon the 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 con-

two tenseless, non-modal sentences: *his method throughout: restate 'if A were, C would be' as a relation between the plain sentences A and C. Here the plain A comes out self-contradictory.*

'counterlegals': *antecedents that conflict with a law, logical or natural. §2's rule will count them all false, since such an antecedent isn't even self-compatible (note 5, p. 12).*

The frame for the rest: the consequent is inferred from the antecedent (1) with relevant conditions and (2) by a law. §2 treats (1), pp. 9–17, §3 treats (2), pp. 17–27.

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

sequent. 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.

the conjunction: *the analysis in outline: A plus true statements S of the conditions yields C. Written A·S from p. 9 on.*

not conditioned upon: *against reading it as 'if the match is dry etc., then if scratched it lights'. The speaker asserts that it's dry. If a condition fails, the counterfactual is false, not left open.*

The principle that permits: *the general statement licensing the step from A·S to C. On p. 17 it's called the connecting principle, and §3 asks which true generalizations can serve as one.*

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.

not a law of logic: *so it can be denied without contradiction, and 'follows by law' isn't entailment. Which general truths count as laws is §3's question (p. 19).*

2. *The Problem of Relevant Conditions*

§2 goes by trial and repair: each definition of S is broken by a case and patched. The patches add up to the rule on p. 13, which the match breaks in its turn on p. 14.

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.

everything follows: *A with not-A is a contradiction, and a contradiction entails any sentence whatever. So C and not-C both follow, and every counterfactual comes out true.*

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 *not-A*—such that from $A \cdot S$ any consequent follows. (Hereafter I shall regularly use the following symbols: "A" for the antecedent; "C" for the consequent; "S" for the set of statements of the relevant conditions or, indifferently, for the conjunction of these statements.)

A·S: *the conjunction of A with the statements in S, the dot being 'and'. Weakening 'all truths' to 'some truths' doesn't help: the one bad truth, not-A, is still available.*

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

a parallel difficulty: *some truths are barred by A through a law of nature, not by logic. Water freezes at 32°F (0°C), so nothing freezes while staying above 33°F.*

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 we have as true generalizations both

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

and also

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.

true generalizations both: *'All Fs are Gs' is true when nothing is F, since nothing refutes it (vacuous truth). No radiator freezes above 32°F, so both hold. With A and S they yield 'breaks' and 'doesn't break'.*

The natural proposal to remedy this difficulty is to rule that counterfactuals cannot 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:

empty laws: *generalizations without instances. The ban is evaded with a disjunct: add soap bubbles and the principle has instances, is true and still covers the radiator. The dimes of p. 23 repeat the trick.*

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 third repair: every member of S must be compatible with A, by logic and by law. Carolina (p. 11) and not-C (p. 12) show how much more than that is needed.

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 cannot 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

Carolina: *no US state has this name. There are two, North Carolina and South Carolina, neighbours on the Atlantic coast. 'In Carolina' means in one or the other.*

all these taken together: *each is compatible with A singly, but together they contradict it. So testing S member by member lets in a set from which anything follows.*

A·S be self-compatible: *now the whole conjunction must be consistent, in logic and by law. But 'for some S' is too easy: one consistent S puts Jones in South Carolina, another in North.*

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

no such set S': *the rule gets a second clause: some S yields C, and no rival set S' of truths yields not-C. Jones now fails both ways, as wanted.*

the negate of the consequent: *C is false, so not-C is a truth. Take it as S': with A it trivially yields not-C. So the second clause defeats nearly every counterfactual, the true ones included.*

testifies: *if A excluded not-C by law, A alone would yield C and no S would be wanted. The match does need its S, so there not-C is compatible with A.*

The diagnosis: S was asked only to deliver C. It must also leave C and not-C open when taken by itself. That gives the rule on p. 13.

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 conditions 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 S (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

[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.

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.

S by itself must not decide: *this bars not-C from S, since not-C isn't compatible with C. Whether C holds must come from adding A to S, not from S.*

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 $-C$.[7] As thus stated, the rule involves a certain redundancy; but simplification is not in point here, for the criterion is still inadequate.

Our rule: *the sum of the repairs since p. 9. Check each clause on its case: 'self-compatible' on the radiator and Carolina, 'no S" on Jones, 'compatible with C and not-C' on not-C.*

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*

The decisive case: a truth compatible with A that wouldn't be true if A were. With it in S, falsehoods get proved. So S must be 'cotenable' with A (p. 15).

---

[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 $A \cdot S$ cannot lead by law to one but not the other. Hence no sentence incompatible with $-A$ can satisfy the other requirements for a suitable S.

incompatible with –A: *then S entails A, and A·S says no more than S. So whatever A·S leads to, S leads to alone, and S would decide between C and not-C, which is already barred.*

[7] Since the first edition of this book, W. T. Parry has pointed out that no counterfactual satisfies this formula; for one can always take $-(A \cdot -C)$ as S, and take $-(A \cdot C)$ as S'. Thus we must add the requirement that neither S nor S' follows by law from $-A$. Of course this does not alleviate the further difficulties explained in the following paragraphs of the text above. (See Parry's 'Reexamination of the Problem of Counterfactual Conditionals', *Journal of Philosophy*, vol. 54 [1957], pp. 85–94, and my 'Parry on Counterfactuals', same journal, same volume, pp. 442–5.)

no counterfactual satisfies: *take –(A·–C), the material 'if A then C', true since A is false. With A it gives C, and –(A·C) gives not-C. So the 'no S" clause never holds unless such S are barred.*

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,

but deny

(ii): *plainly false: scratching wouldn't have wet the match. The test is whether the rule of p. 13 can tell (ii) from (i). The next lines show that it can't.*

(ii) If match *m* had been scratched, it would not have been dry.[8]

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: *true, since m was never scratched. It's compatible with A, as scratched matches can fail to light, and it doesn't settle 'dry' or 'not dry' alone. So the rule lets it into S.*

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·S* we may have

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

a legitimate general law: *the match law of p. 8 read backwards: scratched, well-made, dry matches in oxygen light, so one that is scratched, well made, in oxygen and doesn't light isn't dry.*

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

It was not dry.

no suitable set of sentences S': *'m was dry' is true and would do the job, but S' mustn't settle the consequent alone. And nothing else true, joined to A, gives 'm was dry'.*

And there would seem to be no suitable set of sentences *S'* such that *A·S'* leads by law to the negate of this conse-

[8] 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.

quent. 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*·*S* self-cotenable, if it is not the case that *S* would not be true if *A* were.[9]

a true statement: *'m did not light'. Had m been scratched it would have lit, so this truth doesn't survive the supposition, and it can't be a premise.*

not merely compatible: *cotenable: it's not the case that S would be false if A were true. Itself a counterfactual, hence the regress of p. 16.*

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:

An aside on cotenability in use: which truths stay fixed under a supposition is often left to cues like word order. P. 16 applies this to the Caesar pair of p. 6.

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.

in Georgia: *the US state in the South, not the country. New York City lies in the north-east, about 1,200 km away. So either the city goes south or Georgia reaches north.*

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

---

[9] The double negative cannot be eliminated here; for ". . . if *S* would be true if *A* were" actually constitutes a stronger requirement. As we noted earlier (Note I.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.

The double negative: *'not: S would fail' is weaker than 'S would hold'. The two conditionals are contraries (note 2, p. 6, his 'Note I.2'). Where both are false, S still counts as cotenable.*

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.

Where §2 ends: the relevant conditions of p. 8 can't be defined without counterfactuals. P. 19 adds that cotenability also waits on law, and p. 27 leaves it open.

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

given S is cotenable with A: *for the match: to use 'm was dry' as S we must first know that (ii) of p. 14, 'if scratched, m would not have been dry', is false. That was what had to be explained.*

an infinite regressus or a circle: *regress: each counterfactual waits on another, about S, then S1, and so on without end. Circle: counterfactuals are defined by cotenability, and cotenability by a counterfactual.*

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

no conditions other than the antecedent: *the one way out he considers: begin where A alone leads by law to C, use those cases to test cotenability, and build up. But even the match needs dryness and oxygen.*

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

3. *The Problem of Law*

§3 asks which general statements may carry the step from A·S to C. Being true and general isn't enough, as the coins show on p. 18. So 'law' needs a definition.

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 sustaining a counterfactual; and it is convenient to consider the connecting principles separately.

a matter of logic: *add the law to the premises and C follows by deduction. Nothing is gained: the question becomes which general statements may be added. The two problems aren't cleanly apart (p. 7).*

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: *what you get by putting variables for the names: from 'if m is scratched, dry etc., m lights' to 'every match that is scratched, dry etc. lights'.*

the generalization[10] 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 sustained by the principle thus arrived at, *even* if that principle is *true*. Suppose, for example, that all I had in my right pocket on VE day was a group of silver coins. Now we would not under normal circumstances affirm of a given penny P

VE day: *Victory in Europe Day, 8 May 1945, the end of the Second World War in Europe. Any past day would do. A penny is the US one-cent coin, copper and not silver.*

If P had been in my pocket on VE day, P would have been silver,[11]

even though from

P was in my pocket on VE day

---

[10] 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. 122–43. See also III.3, below.

[11] 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 VE 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 VE 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.

counteridentical: *on that reading P is one of the silver coins and the conditional is true. The example needs the other reading: P, a copper penny, as one more thing in the pocket (see pp. 6, 16).*

we can infer the consequent by means of the general statement

Everything in my pocket on VE 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 sustaining 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.[12]

would not be true: *the supposition overturns the generalization instead of falling under it: with the penny in, not everything there is silver. The match law is left standing by the scratching.*

indeed general, true: *and every case has been checked, so no evidence is lacking. What it lacks isn't support. A clue for p. 20: a law is accepted before its cases are all in.*

thus seems to depend: *the upshot of the coins, and §3's task: to mark off laws from truths that just happen to hold ('casual' means accidental), without using 'would'.*

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. For decision as to the cotenability of two sentences depends partly upon decisions as to whether certain general statements are laws, and we are now concerned directly with

closely related: *to judge whether 'm did not light' would still be true had m been scratched, we use the law that dry, well-made matches light. So cotenability waits on telling laws from accidents.*

[12] 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.

the latter problem. Is there some way of so distinguishing laws from non-laws, among true universal statements of the kind in question, that laws will be the principles that will sustain counterfactual conditionals?

The proposal of §3: a law is a true generalization used to predict cases not yet examined. Two rivals are dismissed first, causal power and logical form.

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

syntactical universality: *being of the form 'Everything that is F is G'. The coin sentence already has it, and any fact about one thing can be given it, so form can't mark off laws.*

All butter melts at 150° F.

from a true and general non-law like

All the coins in my pocket are silver ?

accepted as true: *the mark of a law lies in how we treat it, not in what it says: we rely on it for cases not yet examined. The coin statement is believed only because each coin was looked at.*

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.

the idea behind it: *the bridge to counterfactuals: what we trust for cases not yet examined, we trust for cases that never arise. Prediction can be checked, and lends its authority to them (p. 27).*

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 Humean 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.

Humean idea: *Hume held that we observe no necessity between cause and effect, only regularity and our own expectation. So lawhood doesn't explain predictive use. It consists in it.*

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.

Professor Hempel: *Carl Hempel (1905–97). His papers on confirmation of the 1940s include the raven paradox: does a white shoe, a non-black non-raven, confirm 'all ravens are black'?*

Our criterion excludes vacuous principles as laws. The generalizations needed for sustaining counterfactual conditionals cannot be vacuous, for they must be supported by evidence.[13] The limited scope of our present problem

The empty radiator laws of p. 10, now disposed of: with no instances there's no evidence, so they're never accepted for prediction and can sustain nothing.

[13] 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 $A \cdot S$.

self-compatible: *if A·S broke a law, nothing would satisfy it and its generalization would be empty, so non-vacuity could replace this clause. It can't secure cotenability: 'm did not light' (p. 14) passes it.*

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.

"lawlike": *a law minus its truth: whatever fits a sentence for predictive use, true or not. From here the problem of law is the problem of defining 'lawlike'.*

For convenience, I shall use the term "lawlike" for sentences that, 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.

Two repairs to the definition. Actual use gives way to acceptability (below). Then a contrived law, the dimes of p. 23, forces the final version given there.

Now if we were to leave our definition as it stands, lawlikeness would be a rather accidental and ephemeral property. Only statements that happen actually to have been used for prediction would be lawlike. And 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 nondispositional definition. Before trying to accomplish that, however, we must face another difficulty in our tentative criterion of lawlikeness.

acceptable prior to: *from 'is used' to 'may properly be used': an untried law is lawlike, a fully tested one stays so. But 'acceptable' is dispositional, a hidden 'would' (p. 3). The debt falls due on p. 24.*

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

Everything in my pocket is silver.

All we would need to 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".[14]

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.[15] Naturally this does not mean that

or is a dime: *the dime is the US ten-cent coin, silver in those years. The sentence is true and has unexamined cases, the other dimes, so the p. 22 test makes it lawlike. Compare the soap bubbles, p. 10.*

the pseudo-law: *'P is in my pocket or is a dime' follows by logic from A, and the dime sentence then gives 'P is silver'. So the false penny counterfactual of p. 18 comes out true.*

"dimo": *a made-up word for 'in my pocket or a dime'. It hides the 'or', so banning disjunctions won't help. Compare P on p. 26.*

A sentence is lawlike if: *the final definition. The dime sentence fails it, since we accept it only after checking each thing in the pocket. 'All dimes are silver' passes: no one dime need be checked.*

[14] 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.

[15] 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.

counts vacuous principles as laws: *a vacuous principle has no instances, so its acceptance depends on none, and this test admits it, unlike p. 21's. Read 'class of instances' and the empty class itself must be examined.*

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

no particular instance: *laws still need evidence, but any sample will do: no one case has to be looked at. An accidental generalization is accepted only case by case, with every case counting.*

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

except Number 19: *a generalization with its exception written in. We'd accept it only after looking at marble 19, so it isn't lawlike. Marble d on p. 26 turns the same device against confirmation theory.*

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.

The debt of p. 22 falls due: 'acceptable' must be defined. That's a question for the theory of confirmation, and pp. 25-26 show that the theory has no answer.

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

confirmation: *the support evidence gives a hypothesis. The theories meant (Hempel, Carnap) say roughly that 'All Fs are Gs' is confirmed by Fs found to be G.*

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.[16] Yet obviously in the case of some sentences like

confirmable: *able to gain support from its positive instances. The claim is that instances support some generalizations and not others, however many there are, and no theory says which.*

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,

No twentieth-century president: *the band is about 185 to 187 cm. Even if every president but one fell outside it, we wouldn't bet on the last. Nothing connects the office with height.*

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.

even a few instances: *so the difference isn't in the amount of evidence: many instances don't confirm the pocket sentence, a few confirm the butter one. It lies in the kind of sentence, or predicate (p. 26).*

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

[16] The points discussed in this and the following paragraph have been dealt with a little more fully in my 'A Query on Confirmation', *Journal of Philosophy*, vol. xliii (1946), pp. 383–5.

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 bag by the letters of 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: *a conjunction, each dot being 'and': a is red and b is red, and so on to z. It says that every marble in the bag is red, and 25 of its 26 parts are known to be true.*

R*a*. R*b*. R*c*. R*d*. . . . R*z*

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

–Rd: *the same conjunction with 'd is not red' for 'd is red'. Its other 25 parts are the same known truths, so it has as many favourable cases, and it says the opposite about d.*

R*a*. R*b*. R*c*. R*e*. . . . R*z*. —R*d*,

for again we have 25 favorable and no unfavorable cases. Thus "R*d*" and "—R*d*" 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:

"P": *true of any marble but d if it's red, and of d if it isn't. All 25 examined marbles are P. So 'All the marbles are P' has the same form and the same evidence as 'All the marbles are red'.*

is in the bag 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, and thus 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.

projectible: *a predicate is projectible when its holding of examined cases is a reason to expect it of the rest. 'Red' is, 'P' isn't. Ch. III poses this as the 'grue' riddle, ch. IV answers it.*

So far, I have discovered no way of meeting these diffi-

culties. 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 cannot be directly tested.

Why the gap matters: on p. 20, trust earned by testable predictions was to carry over to counterfactual cases. With acceptability undefined, the authority is unexplained.

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.

Both halves end where p. 3 warned: cotenability is itself counterfactual (p. 16), and law rests on 'acceptable', a disposition term (p. 22) that no theory yet defines.

reduces: *the chain so far: counterfactual to law, law to lawlike, lawlike to acceptable without waiting on any given instance, acceptable to projectible predicates. The last link is missing.*