actually identified as “causes,” both in everyday life and in science, are not strictly sufficient conditions for their alleged effects. For a detailed discussion of problems created by this circumstance for an adequate formulation of an analysis of causal judgments in terms of “constant conjunction,” see my article “Philosophical Analysis, Translation Schemas and the Regularity Theory of Causation,” op. cit., or my book Analytische Erkenntnistheorie, loc. cit.

If we want to express the idea that cause and effect are successive states of the same thing, as in the above example, a more complicated symbolism is needed. But this would not affect the question at issue.

System of Logic, Book III, Chapter V, section 6.

C' may be a negative condition, in which case it would be more natural to speak of the necessary absence of a specified condition. For example, in order for the day-night-day sequence to continue it is necessary that no opaque body interpose itself between sun and earth in such a way that no sunlight could reach any part of the earth's surface at any time.

The following consideration serves as a striking demonstration of the inadequacy of extensional logic for the definition of the concepts “sufficient condition” and “necessary condition” as used in everyday life and science. We often have occasion to say that a certain disjunction of conditions is a necessary condition for a given effect though neither condition by itself is necessary. Thus a college professor may announce to his class that in order to pass his course they must either write a passing term paper or pass a final examination, but that it is not necessary to write a passing term paper, nor is it necessary to pass a final examination. Now, the extensional definition of “q expresses a necessary condition for p” is p ⊃ q. Hence the professor's statement would take the form (p ⊃ q v r) • (~p ⊃ q) • (~p ⊃ r). But since this conjunction entails (p • ~q • ~r), which contradicts the first conjunct, it is self-contradictory.


This seems to be the view W. Kneale takes of laws of nature. As Kneale himself notes in Probability and Induction (Oxford: The Clarendon Press, 1949), p. 71, the view was held by Locke: though man cannot attain to certain knowledge of the laws of nature, as he can only generalize from instances, an angel who knew the “real essences” of natural kinds would see the same sort of necessary connection between causal antecedent and causal consequent as we see between “being a Euclidean triangle” and “being a triangle whose angle-sum equals 180°.”

Of course we may justifiably assert “PA ⊃ QA” even if no instance of P has ever been observed, on the evidence “RA • QA • RB • Qb,” etc., where P and R are similar properties (e.g., let Px = x is water subjected to a temperature exceeding 150°, and Rx = x is coffee subjected to a temperature exceeding 155° and Qx = x boils).

In Burks' system the specified incompatibility, moreover, holds only if the antecedent is physically possible; accordingly there are paradoxes of causal implication analogous to the familiar paradoxes of material and strict implication. This additional proviso, however, has such queer consequences as that “if ice were denser than water, it would not float” and “if ice were denser than water, it would still float” are compatible.

Introduction

(i) Although the following essay attempts to deal in a connected way with a number of connected conceptual tangles, it is by no means monolithic in design. It divides roughly in two, with the first half (Parts I and II) devoted to certain puzzles which have their source in a misunderstanding of the more specific structure of the language in which we describe and explain natural phenomena; while the second half (Parts III and IV) attempts to resolve the more sweeping controversy over the nature of the connection between 'cause' and 'effect,' or, in modern dress, the logical status of 'lawlike statements.'

(ii) The essay begins with a case analysis of a puzzle, taken from recent philosophical literature, relating to the analysis of counterfactual conditionals, statements of the form “If that lump of salt had been put in water, it would have dissolved.” The diagnosis of this puzzle, which occupies the whole of Part I, shows it to rest on a misunderstanding of the conceptual framework in terms of which we speak of what things do when acted upon in certain ways in certain kinds of circumstance. Although the puzzle is initially posed in terms of examples taken from everyday life, the logical features of these examples which, misunderstood, generate the puzzle, are to be found in even the more theoretical levels of the language of science, and the puzzle is as much at home in the one place as in the other. For the framework in which things of various kinds (e.g. matches, white rats) behave ('respond') in various ways (catch fire, leap at a door) when acted upon ('submitted to such and such stimuli') under given conditions (presence of oxygen, 24 hours of food deprivation) is far more basic than the distinctions between metrical and non-metrical concepts, molar and micro-things,
observable and unobservable properties, empirical generalizations and theoretical assumptions, which seem, at first sight, to introduce such a gulf between pre-scientific and scientific discourse.

(iii) If Part I is primarily ‘critical’ in its orientation, calling attention in only the most general terms to the above-mentioned logical features of the framework presupposed by counterfactuals such as “If that match had been scratched, it would have lighted,” and subjunctive conditionals such as “If that piece of salt were put in water, it would dissolve,” Part II attempts a constructive account which, though necessarily brief and schematic, highlights those features of this framework which seem to have caused the most trouble. Postponing for later treatment (Parts III and IV) the classical puzzle about the ‘connection’ between ‘cause’ and ‘effect,’ it explores the logic of expressions for things, for kinds of things, for the causal properties of things, as well as the distinction between properties and states. It offers an analysis of the relation between thing-kinds and the traits in terms of which we identify things as belonging to them which illuminates both the nature and, which is more important, the limitations of the explanations provided by generalizations of the form “Things of kind K behave thusly when such and such is done to them under such and such conditions.” I have italicized the word “limitations” because it is, in my opinion, the considerations advanced at the end of Part II which provide the key to a correct interpretation of the role of theoretical explanations and the status of theoretical (‘unobservable’) entities.

(iv) The second half of the essay (Parts III and IV) is devoted to an attempt to disentangle and resolve the issues matted together in the centuries long debate between the ‘constant conjunction’ (or ‘regularity’) and the ‘entailment’ (or ‘necessary connection’) interpretations of causality. Part III attempts a sympathetic reconstruction of the controversy in the form of a debate between a Mr. C (for Constant Conjunction) and a Mr. E (for Entailment) who develop and qualify their views in such a way as to bring them to the growing edge of the problem. Although it is primarily designed to pose the problem in a way which reflects the philosophical commitments and concerns of the participants in the great debate, Part III also develops some of the themes and distinctions which are put to use in the constructive analysis which follows in Part IV. In particular, it contains a brief discussion of the force of probability statements (section 60), an examination of

what it might mean to say that the world is ‘in principle’ describable without using either prescriptive or modal expressions (sections 79–80), and some remarks on the supposed ‘metalinguistic’ status of modal statements (sections 81–82).

(v) Of the fourth and final part of the essay I shall say only that it offers an account of lawlike statements and of the inductive reasoning by which we support them which shows, in my opinion, how the logical insights of Mr. E can be reconciled with the naturalistic, empiricist tradition defended (if in too narrow and oversimplified a form) by Mr. C.

I. Counterfactuals

1. In his important paper on counterfactual conditionals,* Nelson Goodman interprets his problem as that of “defining the circumstances under which a given counterfactual holds while the opposing counterfactual with the contradictory consequent fails to hold.”† As examples of such opposing counterfactuals, he gives “If that piece of butter had been heated to 150 °F, it would have melted,” and “If that piece of butter had been heated to 150 °F, it would not have melted.”

2. After a quick survey of some varieties of counterfactual and related statements, he finds that “a counterfactual is true if a certain connection obtains between the antecedent and the consequent,”‡ and turns to the task of explaining this connection. He points out, to begin with, that “the consequent [of a counterfactual] seldom follows from the antecedent by logic alone,”§ and never in the case of the empirical counterfactuals with which he is primarily concerned. Nor, in the case of the latter, does the consequent follow from the antecedent alone by virtue of a law of nature. For “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 the conditions are such—i.e., the match is well made, is dry enough,


‡ Ibid., p. 16.

§ Ibid., p. 16.
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 the 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. (p. 17)

"There are," he concludes, "two major problems, though they are not independent and may even be regarded as aspects of a single problem . . . The first . . . is to define relevant conditions: to specify what sentences are meant to be taken in conjunction with the antecedent as a basis for inferring the consequent." * The second is to define what is meant by a law of nature. For 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. (p. 17)

3. Goodman first takes up the problem of relevant conditions. He has implied, in the passages just quoted, that whenever we assert a counterfactual, we have in mind a specific set of relevant conditions, those conditions, indeed, which the relevant law of nature requires to obtain in order that we may infer "That match lights" from "That match is scratched." Instead, however, of focusing attention on these specific conditions, and exploring their bearing on the truth or falsity of the counterfactual, Goodman begins from scratch. Thus he writes,

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. (pp. 17–18)

points out that

if we say that the statement follows by law from the antecedent and

* Ibid., pp. 16–17, passim.

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. (p. 18)

and embarks on the task of so narrowing the class of true auxiliary sentences that we can account for this difference. A compact but lucid argument, in which he introduces a series of restrictions on the membership of this class, leads him to the following tentative rule:

. . . a counterfactual is true if and only if there is some set S of true sentences such that S is compatible with C [the consequent of the counterfactual in question] and with \( \neg C \) [the contradictory consequent], and such that A \( \cdot S \) is self-compatible [A being the antecedent] and leads by law to C; while there is no set \( S' \) compatible with C and with \( \neg C \), and such that A \( \cdot S' \) is self-compatible and leads by law to \( \neg C \). (p. 21)

4. It is at this point that Goodman explodes his bomb.

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, but deny

(ii) If match M had been scratched, it would not have been dry.

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," Goodman continues, without pausing for breath, is caused by including in our S a true statement which although compatible with A would not be true if A were. Accordingly we must ex-
include 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. (pp. 21–22)

5. This new requirement, however, instead of saving the rule leads it to immediate shipwreck.

. . . 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 S₁, 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. (p. 23)

As of 1947, Goodman, “though unwilling to accept this conclusion, [did] not . . . see any way of meeting the difficulty.” * That he still regards this difficulty as genuine, and the line of thought of which it is the culmination philosophically sound, is indicated by the fact that he has made “The Problem of Counterfactual Conditionals” the starting point of his recent re-examination † of the same nexus of problems. Indeed, Goodman explicitly tells us that the four chapters of which this new study consists, and of which the first is a reprinting of the 1947 paper, “represents a consecutive effort of thought on a closely integrated group of problems,” ‡ and that this first chapter contains “an essentially unaltered description of the state of affairs from which the London lectures took their departure.” §

6. It is my purpose in the opening sections of this essay, devoted as it is to fundamentally the same group of problems, to show that Goodman’s puzzle about cotenability arises from a failure to appreciate the force of the verbal form of counterfactuals in actual discourse, and of

the general statements by which we support them; and that this failure stems, as in so many other cases, from too hasty an assimilation of a problematic feature of ordinary discourse to a formalism by reference to which we have succeeded in illuminating certain other features.

7. Let me begin by asking whether it is indeed true that in “the familiar case where for a given match M, we would affirm

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

we would “deny

(ii) If match M had been scratched, it would not have been dry.”

Goodman himself points out in a note that “Of course, some sentences similar to (ii), referring to other matches under special conditions may be true.” * Perhaps he has something like the following case in mind:

Tom: If M had been scratched, it would have been wet.
Dick: Why:
Tom: Well, Harry is over there, and he has a phobia about matches. If he sees anyone scratch a match, he puts it in water.

But just how is Goodman’s “familiar case” different from that of the above dialogue? Why are we so confident that (ii) is false whereas (i) is true? Part of the answer, at least, is that we are taking for granted in our reflections that the only features of the case which are relevant to the truth or falsity of (ii) are such things as that the match was dry, that it was not scratched, that it was well made, that sufficient oxygen was present, that it did not light, “etc.” † and that the generalization to which appeal would properly be made in support of (ii) concerns only such things as being dry, being scratched, being well made, sufficient oxygen being present, lighting, etc. For as soon as we modify the case by supposing Tom to enter and tell us (a) that if M had been scratched, Harry would have found it out, and (b) that if Harry finds out that a match has been scratched, he puts it in water, the feeling that (ii) is obviously false disappears.

8. In asking us to consider this “familiar case,” then, Goodman, whether he realizes it or not, is asking us to imagine ourselves in a

* Ibid., p. 23.
† Fact, Fiction and Forecast. This book, of which the first part is a reprinting of the 1947 paper, contains the University of London Special Lectures in Philosophy for 1953.
§ Ibid., p. 9.
situation in which we are to choose between (i) and (ii) knowing (a) that only the above limited set of considerations are relevant; and (b) that scratching dry, well-made matches in the presence of oxygen, etc. causes them to light. It is, I take it, clear that if we did find ourselves in such a situation, we would indeed accept (i) and reject (ii).

To call attention to all this, however, is not yet to criticize Goodman's argument, though it does give us a better understanding of what is going on. Indeed, it might seem that since we have just admitted that once we are clear about the nature of the case on which we are being asked to reflect, we would, in the imagined circumstances, accept (i) but reject (ii), we are committed to agree with Goodman that the criterion under examination is at fault. For according to it would not (ii) be true?

9. It is not my purpose to defend Goodman's tentative criterion against his criticisms. There are a number of reasons why it won't do as it stands, as will become apparent as we explore the force of counterfactuals in their native habitat. I will, however, in a sense, be defending it against the specific objection raised by Goodman. For it is because he misinterprets the fact that we would accept (i) but reject (ii) that he is led to the idea that the criterion must be enriched with a disastrous requirement of cotenability. And once this fact is properly interpreted, it will become clear that while there is something to Goodman's idea that a sound criterion must include a requirement of cotenability, this requirement turns out to be quite harmless, to be quite free of regress or paradox.

10. But is it, on second thought, so obvious that even if we were in the circumstances described above, we would reject (ii)? After all, knowing that M didn't light, but was well made, that sufficient oxygen was present, etc. and knowing that M wasn't scratched but was dry, would we not be entitled to say,

(iii) If it had been true that M was scratched, it would also have been true that M was not dry?

The fact that this looks as though it might be a long-winded version of (ii) gives us pause.

If, however, we are willing to consider the possibility that (ii) is after all true, the reasoning by which Goodman seeks to establish that if the tentative criterion were sound, (ii) as used in our "familiar case" would be true, becomes of greater interest. The core of this reasoning is the following sentence,

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.

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

It was not dry. (pp. 21-22)

But although Goodman assures us that there is a "legitimate general law" which permits this inference, he does not take time to formulate it, and once we notice this, we also notice that he has nowhere taken time to formulate the "legitimate general law" which authorizes (i).

The closest he comes to doing this is in the introductory section of the paper, where he writes,

When we say

If that match had been scratched, it would have lighted

we mean that the 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." (p. 17)

11. Now the idea behind the above sentence seems to be that the relevant law pertaining to matches has the form

(x)(t) x is a match ⋅ x is dry at t ⋅ x is scratched at t ⋅ implies ⋅ x lights at t

(where, to simplify our formulations, the conditions under which matches light when scratched have been boiled down to being dry.)
And it must indeed be admitted that if this were the "legitimate general law" which authorizes

If M had been scratched, it would have lighted

given M was dry and M was not scratched, there would be reason to expect the equivalent "legitimate general law"

(x)(t) x is a match ⋅ x does not light at t ⋅ x is scratched at t ⋅ implies ⋅ x is not dry at t

to authorize

If M had been scratched, it would not have been dry

given M did not light and M was not scratched.
Or, to make the same point from a slightly different direction, if we
were to persuade ourselves that the laws which stand behind true counterfactuals of the form,

If \( x \) had been \ldots it would have \ldots

are of the form,

\[(x)(t) A(x,t) \cdot B(x,t) \supset C(x,t)\]

we would, in all consistency, expect the equivalent laws

\[(x)(t) A(x,t) \cdot \sim C(x,t) \supset \sim B(x,t)\]
\[(x)(t) B(x,t) \cdot \sim C(x,t) \supset \sim A(x,t)\]

to authorize counterfactuals of the same form. And if we were to persuade ourselves that

[Given that \( M \) was dry, then, although \( M \) was not scratched,] if \( M \) had been scratched, it would have lighted

has the form

[Given that \( x \) was \( A \) at \( t \), then, although \( x \) was not \( B \) at \( t \)] if \( x \) had been \( B \) at \( t \), \( x \) would have been \( C \) at \( t \)

we would expect the equivalent general laws to authorize such counterfactuals as

[Given that \( M \) did not light, then, although \( M \) was not scratched,] if \( M \) had been scratched, it would not have been dry

and

[Given that \( M \) did not light, then, although \( M \) was not dry,] if \( M \) had been dry, it would not have been scratched.

12. But as soon as we take a good look at these counterfactuals, we see that something is wrong. For in spite of the fact that “\( M \) was not dry” can be inferred from “\( M \) was scratched” together with “\( M \) did not light,” we most certainly would not agree that—given \( M \) did not light and \( M \) was not scratched—

If \( M \) had been scratched, it would not have been dry.

What is wrong? One line of thought, the line which leads to cotenability takes Goodman’s “familiar case” as its paradigm and, after pointing out that we are clearly entitled to say

(1) [Since \( M \) was dry,] if \( M \) had been scratched, it would have lighted

but not

(2) [Since \( M \) did not light,] if \( M \) had been scratched, it would not be dry

continues somewhat as follows:

(1) is true
(2) is false

According to (1) \( M \) would have lighted if it had been scratched. But (2) presupposes that \( M \) did not light. Thus (1) being true, a presupposition of (2) would not have been true if \( M \) had been scratched.

So, (1) being true, a presupposition of (2) would not have been true if its ‘antecedent’ had been true. Consequently, (1) being true, the truth of (2) is incompatible with the truth of its ‘antecedent’—surely a terrible thing to say about any conditional, even a counterfactual one . . .

and concludes, (A) that the falsity of (2) follows from the truth of (1); (B) in knowing that (1) itself is true we must be knowing that there is no true counterfactual according to which a presupposition of (1) would not have been true if its [the new counterfactual’s] antecedent had been true; and, in general (C) that in order to know whether any counterfactual, \( \Gamma \), is true we need to know that there is no true counterfactual, \( \Gamma' \), which specifies that the \( S \) required by \( \Gamma' \) would not have been the case if \( \Gamma' \)'s antecedent had been true. In Goodman’s words, “. . . 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 \( S' \), cotenable with \( A \), that leads to \( \sim S \) and so on. Thus we find ourselves involved in a regressus or a circle . . .”

13. Now there are, to say the least, some highly dubious steps in the reasoning delineated above. I do not, however, propose to examine it, but rather to undercut it by correctly locating the elements of truth it contains. That there is something to the above reasoning is clear. The truth of (1) does seem to be incompatible with the truth of (2); and the falsity of (2) does seem to rest on the fact that if \( M \) had been scratched, it would have lighted.

Perhaps the best way of separating out the sound core of the above reasoning is to note what happens if, instead of exploring the logical relationship between the two counterfactuals (1) and (2), we turn
our attention instead to corresponding subjunctive conditionals not contrary to fact in a new “familiar case” which differs from Goodman’s in that these subjunctive conditionals rather than counterfactuals are appropriate. Specifically, I want to consider the “mixed” subjunctive conditionals,

(1’) If M is dry, then if M were scratched, it would light, and

(2’) If M does not light, then if M were scratched, it would not be dry.

Is it not clear as in Goodman’s case that (1’) is true but (2’) false? Indeed, that the falsity of (2’) is a consequence of the truth of (1’)? Here, however, there is no temptation to say that (2’) is false for the reason that in order for it to be true a state of affairs would have to obtain which would not obtain if M were scratched. For (2’), unlike (2) does not require as a necessary condition of its truth that M does not light.

14. How, then, is the incompatibility of (2’) with (1’) to be understood? The answer is really very simple, and to get it, it is only necessary to ask ‘Why would we reject (2’)?’ For to this question the answer is simply that it is just not the case that by scratching dry matches we cause them, provided they do not light, to become wet. And how do we know this? Part of the answer, of course, is the absence of favorable evidence for this generalization; not to say the existence of substantial evidence against it. But more directly relevant to our philosophical puzzle is the fact that in our “familiar case” we are granted to know that scratching dry matches causes them to light. And if this generalization is true—and it must be remembered that we are using “x is dry” to stand for “x is dry and x is well made, and sufficient oxygen is present, etc.”—then the other generalization can’t be true. The two generalizations are, in a very simple sense, incompatible. For if scratching dry matches causes them to light, then the expression ‘scratching dry matches which do not light’ describes a kind of situation which cannot (physically) obtain. And we begin to suspect that Goodman’s requirement of cotenability mislocates the sound idea that (to use a notation which, whatever its shortcomings in other respects, is adequate for the purpose of making this point) if it is a law that

\[(x) Ax \cdot Sx \cdot \supset Cx\]

then it can’t—logically can’t—be a law that

\[(x) Ax \supset \sim Sx\]

15. But we have not yet pinpointed Goodman’s mistake. To do so we must take a closer look at our reasons for rejecting (2’). We said above that we would reject it simply because it is not the case that by scratching dry matches we cause them to become wet. Perhaps the best way of beginning our finer grained analysis is by making a point about our two subjunctive conditionals (1’) and (2’) which parallels a point which was made earlier about counterfactuals (1) and (2).

Suppose that the “legitimate general law” which authorizes (1’) had the form

\[(x)(t) A(x,t) \cdot S(x,t) \cdot \supset C(x,t);\]

would not (2’) be authorized by

\[(x)(t) A(x,t) \cdot \sim C(x,t) \cdot \supset \sim S(x,t)?\]

and hence—in view of the logical equivalence of these two general implications—be true if (1’) is true? Clearly we must do some thinking about the form of the “legitimate general laws” which authorize subjunctive conditionals of the form

if x were . . . it would . . .

and which stand behind contrary-to-fact conditionals of the form

if x had been . . . it would have . . .

This thinking will consist, essentially, in paying strict attention to the characteristics of subjunctive conditionals, counterfactuals and lawlike statements in their native habitat, rather than to their supposed counterparts in PMese.

16. We pointed out above that if we were asked why we would reject (2’) in the context in which it arose, we would say that it is not the case that scratching dry matches causes them to become wet, if they don’t light. We now note that if (1’) were challenged we would support it by saying that scratching dry matches does cause them to light, or that matches light when scratched, provided they are dry, or, perhaps, that if a dry match is scratched it will light, or something of the sort. Is it proper to represent these statements by the form

\[(x)(t) F(x,t) \cdot G(x,t) \cdot \supset H(x,t)?\]

If we leave aside for the moment the fact that there is something odd
about the expression “x is a match at t,” and focus our attention on the other concepts involved, it does not take much logical imagination to see that while ‘there is no law against’ representing “x is scratched at t” by “Sc(x,t), “x is dry at t” by “D(x,t)”, and “x lights at t” by “L(x,t)”, to do so is to obscure rather than make manifest the logical form of “If a dry match is scratched, it lights.” For it is by no means irrelevant to the logic of this generalization that matches begin to burn when they are scratched. And it is a familiar fact that

As B when Ded—provided the circumstance, C, are propitious, concerns something new that A’s begin to do when changed in a certain respect in certain standing conditions—which need not, of course, be ‘standing still.’

17. I do not, by any means, wish to suggest that all empirical generalizations are of the above form. Clearly,

Eggs stay fresh longer if they are not washed,
which authorizes the counterfactual

If this egg had not been washed, it would have stayed fresh longer,
is not of this form. But our problem, after all, is that of understanding just why it is clear that—in Goodman’s “familiar case”—we would affirm

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

but reject

(ii) If M had been scratched, it would not have been dry,

and to do this we must get the hang of generalizations of the former kind.

18. Now, being dry is obviously not the same thing as becoming dry, nor beginning to burn as burning, and though we can imagine that someone might say “matches burn when scratched,” this would, strictly speaking, be either incorrect or false—incorrect if it was intended to express the familiar truth about matches; false if it was intended to express the idea that matches burn when they are being scratched as iron rusts while exposed to moisture. (Having made this point, I can now rephrase it by saying that if it were correct to use “matches burn when scratched” in the former sense, this would simply mean that “burn” has an idiomatic use in which it is equivalent to “begin to burn”.)

With this in mind, let us examine the apodosis of Goodman’s (ii),

COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES

namely, “. . . it (M) would not have been dry.” If we suppose that this is intended to have the force of “would have become wet”, we can, indeed, assimilate

If M had been scratched, it would not have been dry
to the form

If x had been Ded, it would have Bed

of which (i) is such a straightforward example. For becoming wet would seem to be a legitimate example of B-ing.

But while there are true generalizations to the effect that doing certain things to matches in certain favorable circumstances causes them to become wet, none of them seem to involve scratching. Again, a match which becomes wet must have been dry, and approaching Goodman’s “familiar case”—as we do—in the knowledge that scratching dry (etc.) matches causes them to light, we cannot consistently say both

[Since M was dry, etc.,] if M had been scratched, it would have lighted

and

[Since M was dry, etc.,] if M had been scratched, it would have become wet

unless we suppose that the circumstances in which scratching dry matches causes them to become wet (the etc. of the second ‘since’ clause) differs in at least one respect from the circumstances in which scratching dry matches causes them to light (the etc. of the first ‘since’ clause). And it is clearly no help—in the absence of this supposition—to add to the second counterfactual the proviso, “provided M does not light”; for this proviso, given the truth of the first counterfactual, is physically inconsistent with the conjunction of the antecedent of the second counterfactual with the ‘since’ clause on which it rests.

19. If, therefore, we interpret “. . . it (M) would not have been dry” as “. . . it (M) would have become wet,” we run up against the fact that a generalization is implied which is not only patently false, but inconsistent—given the stipulations of the case—with one which we know to be true. And this is, as we have already noted, the sound core of Goodman’s cotenability requirement. Two counterfactuals cannot both be true if they imply logically inconsistent generalizations. If one counterfactual is true, no counterfactual which involves an ante-
Wilfrid Sellars
cedent-cum-circumstances which is specified to be physically self-incompatible by the generalization implied by the first counterfactual can also be true. On the other hand, cotenability thus understood leads to no “infinite regressus or . . . circle,” for while one has not confirmed a generalization unless one has disconfirmed logically incompatible generalizations, this does not mean that before establishing one thing one must first establish something else, and so on. For the process of confirming a generalization is the process of disconfirming logically incompatible generalizations.

20. Suppose, however—as is indeed obvious—that we are not to interpret “. . . would not have been dry” as “. . . would have become wet”; does another interpretation of (ii) lie within groping distance? The answer is Yes—but, as before, on condition that we are prepared to make certain changes in its wording. Let us begin this groping with an examination of—not (ii) but—the closely related counterfactual,

If M had been scratched without lighting, then it . . .
then it what? Should we say “. . . would not have been dry”? or “. . . could not have been dry”? Clearly the latter. The difference, in this context, between ‘w’ and ‘c’ is all important. It is the difference between

(A) Matches will not be (stay) dry, if they are scratched without lighting

and

(B) Matches cannot be dry, if they do not light when scratched.

(A) introduces, as we have seen, a new generalization into our “familiar case”—one which is inconsistent with

(C) Matches will light when scratched, provided they are dry, which is the generalization implied by (i). (B), on the other hand, far from being inconsistent with (C) would seem to be just another version of it.

And it is clear, on reflection, that (C) is the only ‘will’ statement which expresses the fact that scratching dry matches causes them to light. Thus, we can say that

A dry match will light when scratched

but not

A match which does not light when scratched will not be dry.

COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES

To be sure, we can say—with a little license—

A match which does not light when scratched will be found not to be dry

so that the above claim is not quite true. But the point being made is clear enough, and, in any case, we shall be examining the ‘exception’ in a moment.

21. We begin, therefore, to suspect that corresponding to generalizations of the form

Bing As causes them to D—provided C

there is only one correctly formed counterfactual of the form “If x had been Yed it would have . . . ” namely

[Since C,] if this A had been Bed, it would have Ded

which is not to say that each such generalization might not authorize a number of counterfactuals having a different form. Beating about the bushes for other asymmetries pertaining to our familiar generalization about matches, we notice that while it tells us that scratching matches causes them to light, it doesn’t tell us the cause of matches not being dry; and that while it enables us to explain the fact that a match lighted on a certain occasion by pointing out that it was scratched and was dry, it doesn’t enable us to explain the fact that a match was not dry by pointing out that it was scratched without lighting.

On the other hand, the generalization does enable us to explain how we know that a given match was not dry. “I know that it wasn’t dry, because it didn’t light when scratched.” “M can’t have been dry, because it was scratched, but did not light.” “Since M was scratched, but did not light, it can’t have been dry.” “M was scratched without lighting, so it wasn’t dry.” All these point to the hypothetical

(M was scratched without lighting) implies (M was not dry),

and, indeed, to the general hypothetical

The fact that a match is scratched without lighting implies that it was not dry.

22. I have already pointed out how misleading it is to characterize the “legitimate general law” which authorizes the counterfactual

If M had been scratched, it would have lighted
as a “principle which permits the inference of
That match lights
from
That match is scratched. That match is well made. Enough
oxygen is present. Etc.” *

For the fact that if there is a principle which authorizes the inference
of \( \sim S_2 \) from \( S_1 \cdot \sim S_3 \), there will also be a principle which authorizes the
inference of \( \sim S_2 \) from \( S_1 \cdot \sim S_3 \) leads one to expect that the same
general fact about matches which, in Goodman’s “familiar case,” supports
the above counterfactual, will also support

If M had been scratched, it would not have been dry,
an expectation which is the ultimate source of the puzzlement exploited
by Goodman’s paper.

23. This is not to say that it is wrong to interpret our generalization
about matches as a “season inference ticket.” It is rather that the connection
between the generalization and the counterfactual, “If M had
been scratched, it would have lighted,” rests on features of the
generalization which are not captured by the concept of a season inference
ticket, and which, therefore, the logical form of a general hypothetical
does not illuminate. Thus, while

\[(m)(t) \text{ m is scratched at } t \cdot m \text{ is dry at } t \cdot \text{ implies } \cdot m \text{ lights at } t \dagger\]
does, in a sense, have the force of “dry matches light if scratched,” or
“scratching dry matches causes them to light,” this mode of representation
must be supplemented by a commentary along the lines of the
above analysis, if its relation to “If M had been scratched, it would
have lighted” is to be understood; while if our familiar fact about
matches is assimilated without further ado to the form

\[(x)(t) A(x,t) \cdot B(x,t) \cdot C(x,t) \cdots \supset L(x,t)\]

all chance of clarity has been lost.

24. We have connected the fact that scratching a match is doing
something to it, with the fact that we expect

\[\ldots \text{ if M had been scratched } \ldots\]

* Fact, Fiction and Forecast, p. 17.
† I have used the so-called ‘tenseless present’—a typical philosophical invention—
to simplify the formulation of this general hypothetical, without, in this context,
doing it too much violence.

COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES
to be preceded, at least tacitly, by an expression referring to standing
conditions, thus

\[\text{[Since M was dry,]} \ldots\]
and to be followed by an expression referring to a result, thus

\[\ldots \text{ it would have lighted.}\]

It is important to bear in mind that the distinction between the standing
conditions, the doing and the result is an objective one. It is not
relative to a particular way of formulating the general fact that dry
matches light when scratched. The equivalent formulas

\[(x)(t) D(x,t) \supset \cdot \text{Sc}(x,t) \supset L(x,t)\]

\[(x)(t) \sim L(x,t) \supset \cdot \text{Sc}(x,t) \supset \sim D(x,t)\]

\[(x)(t) \sim L(x,t) \supset \cdot \text{D}(x,t) \supset \sim \text{Sc}(x,t)\]

do not give us three different ways of cutting up the above fact about
matches into a standing condition, a doing and a result or consequence;
although, in a purely logical sense, “\( \sim L(x,t) \)” may be said to formulate
a ‘condition’ under which “\( (x)(t) \text{Sc}(x,t) \supset \sim D(x,t) \)” holds, and
“\( \text{Sc}(x,t) \)” and “\( \sim D(x,t) \)” respectively to be the ‘antecedent’ and the
‘consequent’ of this implication.

The fact that “\( D(x,t) \)” formulates a ‘condition’ in a sense in which
“\( \sim L(x,t) \)” does not, is, though obvious, the key to our problem. For
it is just because “If M were scratched . . .” and “If M had been
scratched . . .” are expressions for somethings being done to something (in a certain kind of circumstance) that we expect them to be followed, not just by a ‘consequent,’ but by (expressions for) a consequence, and also expect the context to make it clear just what conditions or circumstances are being implied to obtain. There is, however,
a manner of formulating this same content which does not evoke these
expectations, and which does focus attention on specifically logical relationships. Consider, for example, the following conditionals,

\[\text{If it were the case that M was scratched without lighting, it}
\text{would be the case that M was not dry.}\]
\[\text{If it had been the case that M was scratched without lighting,}
\text{it would have been the case that M was not dry.}\]

Clearly, to wonder what else would have to be the case, if it were
the case that M was scratched without lighting, is not the same thing
as to wonder what the consequence of striking a match would be, given
that it failed to light.
Wilfrid Sellars

Does this mean that in our familiar case, we would accept
If it had been the case that M was scratched, it would have been the case that M was not dry (or would not have been the case that M was dry)

although we reject
If M had been scratched, it would not have been dry?
The answer is almost Yes. We are getting ‘warmer,’ though there is still work to be done. I shall introduce the next step by discussing examples of quite another sort.

25. Consider the following (where n is a number, perhaps the number of planets):

(1) If n were divisible by 3 and by 4, it would be divisible by 12
(2) If n were divisible by 3, then if n were divisible by 4, it would be divisible by 12
(3) Since n is divisible by 3, if n were divisible by 4, it would be divisible by 12
(4) Since n is divisible by 3, if n had been divisible by 4, it would have been divisible by 12

and then the following (where n is, say, the number of chess pieces on a side):

(5) If n were not divisible by 12, but divisible by 3, it would not be divisible by 4
(6) If n were not divisible by 12, then if n were divisible by 3, it would not be divisible by 4
(7) Since n is not divisible by 12, if n were divisible by 3, it would not be divisible by 4
(8) Since n is not divisible by 12, if n had been divisible by 2, it would not have been divisible by 4.

The crucial step, in each series, is from the first to the second, i.e. from (1) to (2), and from (5) to (6). (1) and (5) are clearly true. What of the others? What, to begin with, shall we say about (2)? The point is a delicate one. At first glance, it looks quite acceptable, a sound inference ticket. But would we not, perhaps, be a bit happier if it read

(2') If n were divisible by 3, then if it were also divisible by 4, it would be divisible by 12?

What of (3)? It calls attention to the argument,
n is divisible by 3
So, if n were divisible by 4, it would be divisible by 12.

COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES

(Call the conclusion of this argument C1.) The principle of the argument is the complex hypothetical (2). And the question arises, How does this argument differ from one that would be authorized by (1)? Consider the argument

n is divisible by 3
So, if it were also divisible by 4, it would be divisible by 12.

There is clearly a sense in which the conclusion of this argument (call it C2) is more cautious than that of the preceding argument. For C2 carries with it a reference to, and a commitment to the truth of, its premise, which is lacking in C1. The difference can be put by noting that while C1, as a conclusion, may imply that one has come to know that ‘n is divisible by 12’ can be inferred from ‘n is divisible by 4’ by virtue of knowing that n is divisible by 3; it does not imply that what one has come to know is ‘that ‘n is divisible by 12’ can be inferred from ‘n is divisible by 4’ given that n is divisible by 3’. And by not implying this, indeed by implying that what one has come to know is ‘that ‘n is divisible by 12’ can without qualification be inferred from ‘n is divisible by 4’; it is false. To infer from p the legitimacy of the inference from q to r, is not the same thing at all as to infer from p the legitimacy of the inference from q to r, given p.

26. In the symbolism of modal logic, there is all the difference in the world between ‘p  q  r’ and ‘p  (q  r)’ even though the corresponding formulas in the system of material implication are equivalent. The former authorizes the argument

I. p
   So, q  r

but not, as does the latter

II. p
   So, q  r

And while argument II defends the subjunctive conditional

If q were the case, r would be the case

argument I does not. The only resembling subjunctive conditional defended by the assertion that p, and an appeal to ‘p  q  r’ is

If q were the case as well as p, r would be the case
and while this latter carries with it the assertion that \( p \), it does not do so in the same way as does

[Since \( p \) is the case,] if \( q \) were the case, \( r \) would be the case.

For this points to argument II with its stronger conclusion.

27. Turning back to our two lists of conditional statements about numbers, we can now see that the counterfactual corresponding to (1) is not (4), but rather

(4') Since \( n \) is divisible by 3, if it had also been divisible by 4, it would have been divisible by 12.

On the other hand, (4) does correspond to (2). Again, (6) has the same force as (5) only if it is interpreted as

(6') If \( n \) were not divisible by 12, then if it were also the case that it was divisible by 3, it would not be divisible by 4.

and (5) authorizes (8) only if it is interpreted as

(8') Since \( n \) is not divisible by 12, if it had also been the case that it was divisible by 3, it would not have been divisible by 4.

28. It is worth noting, in this connection, that it is not only in cases like our match example, where we are dealing with particular matters of fact, that a generalization may enable us to explain how we know a fact, without enabling us to explain the fact itself. Thus, while

If \( n \) is divisible by 3 and 4, it is divisible by 12

enables us to explain the fact that a certain number is divisible by 12 (“It is divisible by 12 because it is divisible by 3 and by 4.”), it does not enable us to explain the fact that a certain number is not divisible by 4, though it does enable us to explain how we happen to know that the number is not divisible by 4. (“I know that it is not divisible by 4, because though it is divisible by 3, it is not divisible by 12.” “He knows that it is not divisible by 4, because he knows that though it is divisible by 3, it is not divisible by 12.” “It can’t be divisible by 4, because, though it is divisible by 3, it is not divisible by 12.”) It would simply be a mistake to say, “It is not divisible by 4, because, though divisible by 3, it is not divisible by 12.”

One is tempted to put this by saying that just as one explains a particular matter of empirical fact by ‘showing how it comes about,’ and not, simply, by subsuming it under the ‘consequent’ of a general hypothetical, the ‘antecedent’ of which it is known to satisfy, so one explains such a fact as that a certain number is divisible by 12, or not divisible by 4, not simply by subsuming it under the ‘consequent’ of any old mathematical truth under the ‘consequent’ of which it can be subsumed, but only by applying a mathematical truth which, so to speak, takes us in the direction of the ‘genesis’ of the property in question in the mathematical order, i.e. which starts us down (or up?) the path of what, in a neatly formalized system, would be its ‘definition chain.’ Certainly “\( n \) is divisible by 12, because it is divisible by 3 and by 4” has something of this flavor, while “\( n \) is not divisible by 4, because, though divisible by 3, it is not divisible by 12” does not. But to say anything worthwhile on this topic, one would have to say a great deal more than there is space for on this occasion.

29. Now the moral of these mathematical examples is that the counterfactual most resembling Goodman’s (ii) which is authorized by our (simplified) generalization about matches, is, explicitly formulated,

(ii') Since \( M \) did not light, if it had also been the case that it was scratched, it would have been the case that it was not dry,

and it would be correct to boil this counterfactual down to

If it had been the case that \( M \) was scratched, it would have been the case that \( M \) was not dry

only if the context makes it clear that there is a tacit also in the statement, and indicates in what direction the additional presupposition is to be found.

Why, then, it may be asked, should we not conclude that (i) itself is simply a shorter version, appropriately used in certain contexts, of

(i') Since \( M \) was dry, if it had also been the case that \( M \) was scratched, it would have been the case that \( M \) lighted?

(Clearly it is not a shorter version of

Since \( M \) was dry, if it had also been scratched . . .

for this would imply that something else must be done to the match besides scratching it, to make it light.)

The answer should, by now, be obvious. It is part of the logic of generalizations of the form

\[ X' \text{ causes } Y' \text{ to } Z', \text{ provided . . .} \]
that when we say

If this Y had been Xed, it would have Zed,
it is understood that it is because this Y was in certain (in principle)
specifiable circumstances that it would have Zed if it had been Xed. In
other words, the fact that it is proper to say, simply
If M had been scratched, it would have lighted
rests on the relation of the statement to the objective distinction be-
tween standing conditions, what is done and its result.

30. If, however, the framework of our discussion has been adequate
for the purpose of dispelling the specific perplexities generated by Good-
man’s formulation of his first problem, we must, before we turn our
attention to his second problem, namely, that of the ‘connection’ be-
tween the antecedent and the consequent of a law of nature construed
as a general hypothetical, build this framework into some sort of over-
all constructive account of the logical form of what, for the time being,
we shall lump together as ‘causal generalizations in actual usage.’

II. Thing-Kinds and Causal Properties

31. I shall begin this constructive account of causal generalizations
with some remarks which grow quite naturally out of the first part of
this essay. Suppose we have reason to believe that

Φ-ing Ks (in circumstances C) causes them to ψ

(where K is a kind of thing—e.g., match). Then we have reason to be-
lieve of a particular thing of kind K, call it x₁, which is in C, that

x₁ would ψ, if it were Φ-ed.

And if it were Φ-ed and did ψ, and we were asked “Why did it ψ?” we
would answer, “Because it was Φ-ed”; and if we were then asked, “Why
did it ψ when Φ-ed?” we would answer “Because it is a K.” If it were
then pointed out that Ks don’t always Φ when ψ-ed, we should counter
with “They do if they are in C, as this one was.”

31. Now, there is clearly a close connection between

x₁ is (water-) soluble

and

If x₁ were put in water, it would dissolve.

So much so, that it is tempting to claim, at least as a first approximation,

that statements of these two forms have the same sense. I believe that
this claim, or something like it, would stand up under examination—
indeed, that the prima facie case in its favor is so strong that to defend
it is simply to weed away misunderstandings. Unfortunately, “simply”
to weed away misunderstandings is not a simple job. For most of them
spring from misguided efforts to fit causal discourse into an overly
austere, indeed procrustean, empiricism. And it will not be until the
conclusion of the fourth and final section of this essay—in which I shall
attempt to clarify certain fundamental issues pertaining to scientific
inference—that we will, I believe, be in a position to accept the ‘ob-
vious’ with good philosophical conscience.

32. Perhaps the simplest way to come to grips with puzzles about
causal properties is to represent the analysis of concepts like ‘(water-)
soluble’ by the schema

D(x,t) = df Φ(x,t) implies ψ(x,t)

(where ‘Φ(x,t)’ and ‘ψ(x,t)’ are informally construed as the counter-
parts, respectively, of ‘x is scratched at t’ and ‘x lights at t’) and then
ask “What is the force of the term ‘implies’ in this context?” For this
question calls attention to the fact that to attribute a property of the
sort we are considering to an object is to be prepared in that context
to infer ‘it ψs’ from ‘it is Φ-ed.’ And if the phrase “in that context” poses,
in a sense, our original problem all over again, there may be some gain
in the reformulation.

Now, any answer to this question must account for the fact that we
think of being Φ-ed as the cause of ψ-ing. The pellet of salt dissolves
because it is put in water, just as, in our earlier example, the match
lights because it is scratched. This suggests that to attribute to an object
a ‘disposition’ of which the ‘antecedent’ is being Φ-ed and the ‘conse-
quent’ ψ-ing, is to commit oneself to the idea that there is a general
causal fact, a law, shall we say, which relates being Φ-ed to ψ-ing.

On the other hand, it is perfectly clear that this law is not to the
effect that Φ-ing anything causes it to ψ, i.e. that

(x)(t) Φ(x,t) implies ψ(x,t)

For when we say of a piece of salt that it is soluble, we are certainly
not committing ourselves to the idea that everything, e.g. a stone, dis-
solves in water. And, indeed, the argument of the first part of this essay
has made it clear that if, when we attribute a ‘disposition’ to an object,
we are committing ourselves to the existence of a general fact involving \( \psi \)-ing and being \( \Phi \)-ed, this general fact has the form \( \Phi \)-ing \( Ks \) (in \( C \)) causes them to \( \psi \).

And reflection on the fact that when we attribute a ‘disposition’ to an object we think of being \( \Phi \)-ed as the cause of \( \psi \)-ing calls attention to the fact that words like “dissolves”, “ignites”, etc. are words for results. To say of something that it has dissolved is to say more than that, having been placed in water, it has disintegrated and disappeared, but is recoverable, say, by evaporation. It is to imply that it has disintegrated because it was placed in water. It is no accident that alongside such a word as “soluble” we find the word “dissolves”. And this, in turn, suggests that it is no accident that alongside such a word as “soluble” there is the fact that we know such general truths as that salt dissolves in water.

33. Let us, therefore, work for the time being with the idea that when we ascribe a ‘disposition’ to a thing, we are committing ourselves to the idea that there is a general fact of the form

Whenever and wherever a thing of kind \( K \) is \( \Phi \)-ed (in favorable circumstances) it \( \psi \).

And let us limit our discussion to those cases in which the things referred to are correctly classified by a thing-kind word in actual usage, and in which the causal properties in question are similarly enshrined in discourse. In other words, let us examine the logic of ‘disposition terms’ in a framework which abstracts from the fact, so annoying to logicians, that human discourse is discourse for finding things out as well as for expressing, in textbook style, what we already know.

34. Let us suppose, then, that to ascribe the causal property \( D(x,t) \) to \( x_1 \) now, where (it is also supposed that)

\[
D(x,t) = \text{If } \Phi(x,t) \text{ implies } \psi(x,t)
\]

is to commit oneself to the idea that \( x_1 \) belongs to a kind of thing, \( K \), and is in a kind of circumstance, \( C \), such that if one knew which kind was \( K \), and which kind was \( C \), one would be in a position to reason

\[
x_1 \text{ is } K, \text{ and is in } C
\]

So, if it were also the case that \( \Phi(x_1, \text{now}) \), it would be the case that \( \psi(x_1, \text{now}) \)

COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES

thus,

\[
x_1 \text{ is a pellet of salt} \ldots
\]

So, if \( x_1 \) were put in water, it would dissolve.

Now, the “also” in the above reasoning schema reminds us that the first part of this essay has made it clear that

If \( x \) were \( \Phi \)-ed, it would \( \psi \)

is correctly transcribed into the technical language of logic by neither an unqualified

\[
\Phi(x_1, \text{now}) \supset \psi(x_1, \text{now})
\]

which is obvious, nor by

\[
\Phi(x_1, \text{now}) \rightarrow \psi(x_1, \text{now})
\]

where ‘\( \rightarrow \)’ is the basic symbol for ‘causal implication.’ The root idea behind modal connectives is inferability, and, as we saw, once we turn from ordinary discourse to logical formulations, the above subjunctive conditional must, in the first instance (i.e. neglecting, for the moment, all the other respects in which such formulations are misleading), be represented by the schema

\[
\ldots \text{ if it were also the case that } \Phi(x_1, \text{now}), \text{ then it would be the case that } \psi(x_1, \text{now})
\]

where the also makes it clear that the above attempt to represent the conditional by an unqualified modal statement simply won’t do.

35. On the other hand, our analysis has also made it clear that the force of a subjunctive conditional is, at bottom, a modal force. It rests on such inference-authorizing general truths as ‘salt dissolves in water’ (or, to mention an example of a kind which is excluded from the restricted framework of the present discussion, ‘(In the northern hemisphere) floating needles point to the northernmost regions of the earth’) And though we shall not come to grips with the ‘causal modalities’ until the concluding section of this essay, it is not unreasonable to assume, provisionally, that these general truths are properly represented, in the technical language of modal logic, by the form

\[
(x)(t) : K(x) \cdot C(x,t) \rightarrow \Phi(x,t) \supset \psi(x,t)
\]

If so, it would be natural to suggest that

If \( x_1 \) were \( \Phi \)-ed, it would \( \psi \)

as presupposing a general fact of this form, should be transcribed by
Wilfrid Sellars

an implication symbol which has a modal force without being the basic symbol for causal implication. It might, for example, be represented by the use of a shorter arrow, thus
\[ \Phi(x_1, \text{now}) \rightarrow \psi(x_1, \text{now}). \]

36. But before we follow up this suggestion, let me call attention to the fact, which has undoubtedly been noticed, that the above general hypothetical contains the expression ‘\( K(x) \)’, and not, as might have been expected, the expression ‘\( k(x,t) \)’. This formulation embodies the fact that where ‘\( K \)’ is a thing-kind word, it is misleading to represent both ‘\( x \) is a \( K \)’ and, say, ‘\( x \) is red’ by the same form, ‘\( F(x,t) \)’. The point is simply that being a \( K \) is logically related to the self-identity of a thing, \( x \), at different times, in a way in which being red is not. It might be put by saying that where ‘\( K \)’ is a thing-kind word in a given context of discourse, if at any time it is true of \( x \) that it is a \( K \), then if \( x \) were to cease to be a \( K \), it would cease to be \( x \), i.e. would cease to be (full stop). Or, to put it less paradoxically, being a \( K \) is not something that a thing is at a time, though it may be true at a time that it is a \( K \). “Can’t we say that \( x \) was a child at \( t \), and subsequently a man at \( t’ \)?” But ‘child’, as Aristotle saw, is not a thing-kind term in the sense in which ‘man’ (human being) is a thing-kind term. Child and grown-up are not sub-kinds of man (human being), as man and dog are sub-kinds of animal.

How we classify objects depends on our purposes, but within a given context of discourse, the identity of the things we are talking about, their coming into being and ceasing to be, is relative to the kinds of that context.

37. I shall have more to say about thing-kind words in a moment. But first let us put the suggestion of section 36 to work in the analysis of ‘disposition terms.’ With the introduction of the symbol ‘\( \rightarrow \)’ we would seem to be in a position to answer the question “What kind of implication belongs in the schema
\[ D(x,t) \rightleftharpoons \Phi(x,t) \implies \psi(x,t)? \]
by simply rewriting it as
\[ D(x,t) \equiv \Phi(x,t) \rightarrow \psi(x,t). \]
But if we do so, we must not forget that it is only because we have informally been construing ‘\( \Phi(x,t) \)’ as ‘\( x \) is \( \Phi \)-ed at \( t \)’ and ‘\( \psi(x,t) \)’ as

\[ \text{COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES} \]

‘\( x \) \( \psi \)s at \( t \)’ that we can go from
\[ \Phi(x_1, \text{now}) \rightarrow \psi(x_1, \text{now}) \]
to
If \( x_1 \) were \( \Phi \)-ed, it would \( \psi \)
as contrasted with
If it were also the case that \( \Phi(x_1, \text{now}) \), it would be the case that \( \psi(x_1, \text{now}) \).

In other words, the informal commentary with which we have been surrounding our use of logistical expressions is essential to their correct interpretation as a transcription of causal discourse. This commentary is associated with a division of the functions which appear in this transcription into four categories with a different sign design for each category (thus, ‘\( K_1 \)’ ‘\( K_2 \)’ . . . ‘\( K_n \); ‘\( C_1 \)’ ‘\( C_2 \)’ . . . ‘\( C_n \); ‘\( \Phi_1 \)’ ‘\( \Phi_2 \)’ . . . ‘\( \Phi_n \); ‘\( \psi_1 \)’ ‘\( \psi_2 \)’ . . . ‘\( \psi_n \)’) such that expressions for kinds of things are transcribed by a ‘\( K \)’, expressions for kinds of circumstances by a ‘\( C \)’, expressions—roughly—for something done to a thing by a ‘\( \Phi \)’ and expressions for what it does in return by a ‘\( \psi \)’. Thus, the form ‘\( F(x,t) \)’ has a different logic depending on whether ‘\( F \)’ is representing a ‘\( C \)’, a ‘\( \Phi \)’, or a ‘\( \psi \)’. And, indeed, if we read ‘\( F(x,t) \)’ as ‘\( x \) is \( F \) at \( t \),’ we should not be under the illusion that “at \( t \)” means the same whether ‘\( F \)’ is a ‘\( C \)’, a ‘\( \Phi \)’, or a ‘\( \psi \);’ or, for that matter—if we were to elect to represent ‘\( x \) is a \( K \)’ by ‘\( K(x,t) \)’—a ‘\( K \)’.

Thus it will not do simply to propose
\[ D(x,t) \equiv \Phi(x,t) \rightarrow \psi(x,t) \]
as the technical transcription of
\[ x \text{ is soluble at } t \text{ if and only if, if } x \text{ were placed in water at } t, \text{ } x \text{ would (begin to) dissolve at } t \]
stipulating only that expressions of the form
\[ f_1(x_1,t_1) \rightarrow f_2(x_1,t_1) \]
implicate the truth of a statement of the form
\[ (x)(t) f_3(x,t) \cdot f_1(x,t) \rightarrow f_2(x,t) \]
and the truth of
\[ f_3(x_1,t_1). \]
For, in the absence of such additional stipulations as have been asso-
associated with our division of descriptive functions into four categories, the above transcription procedures would generate Goodman's paradox.

38. Now we could, of course, abandon the attempt to capture the force of such subjunctives as

- If M were scratched, it would light

in our technical language, and limit ourselves instead to conditionals of the form

- If p were the case, q would be the case,

abandoning the device of contextual implication, and putting all implications into the direct content of what is said. This would mean that the counterpart of

- If this were put in water, it would dissolve

would no longer be

\( \Phi(x_1, \text{now}) \rightarrow \psi(x_1, \text{now}) \)

but rather something like

\[
(\exists K)(\exists C) : K(x_1) \cdot C(x_1, t_1) : (x)(t) : K(x, t) \cdot C(x, t) \rightarrow \cdot \\
\Phi(x, t) \supset \psi(x, t)
\]

i.e., there is a kind of thing and a circumstance such that, if it is of that kind and is now in that circumstance, it is such that if anything of that kind is ever in that circumstance, then it is \( \Phi \)-ed, it \( \psi \)-s. For the only way in which the contextual implications of the above subjunctive conditional could become part of the direct content of what is asserted, is by the use of existential qualification over thing-kind and circumstance variables.*

On the other hand, if our technical language does distinguish between the four classes of function, there is no reason why we should not introduce a symbol, ‘”‘, together with the contextual stipulations described above. The issue is partly a matter of what we want our technical language to do. If our aim is the limited one of rewriting ordinary causal property and subjunctive conditional discourse in a symbolism sprinkled with ‘\( \psi \)’s, ‘\( \psi \)’s, etc., then this purpose is readily achieved by introducing as many categories of, and logics for, symbolic expressions as are necessary to reproduce the complexity of ordinary usage. If, on

the other hand, our aim is in some sense to analyze or reconstruct ordinary usage, then, instead of simply creating, so to speak, a symbolic code for ordinary causal discourse, we will seek to introduce these special categories of expression with their special ‘logics,’ in terms of a smaller number of initial categories and a basic framework of logical principles. This effort, presumably, would be guided, not so much by abstract considerations of formal elegance, as by reflection on the scientific use the product is to have. Of course, once the appropriate derivative categories had been introduced, it would then be possible to introduce the symbol ‘”‘’ as before. This time, however, our transcriptions of causal discourse would be far more than a simple rewriting in logistical symbols.

39. Now it might be thought that the task of constructing our four categories of function, K, C, \( \Phi \), and \( \psi \) out of more primitive descriptive functions, including, among others, a function ‘\( f(x, t) \)’ (e.g. ‘Red(x, t)’), is a straightforward one, if not downright easy. That this is not the case; that this task is not only not easy, but that it may spring from a misconception, will emerge, I believe, in the following paragraphs.

40. We pointed out above that when one makes explicit the presuppositions of a statement to the effect that a certain object, \( x \), has a certain causal property \( D \) (of the kind we are considering) at time \( t \), thus, ‘\( D(x_1, t_1) \)’, one gets something like

\[
(\exists K)(\exists C) : K(x_1) \cdot C(x_1, t_1) : (x)(t) : K(x, t) \cdot C(x, t) \rightarrow \cdot \\
\Phi(x, t) \supset \psi(x, t)
\]

Bearing in mind our stipulations of an ideal universe of discourse containing a fixed and known and named variety of thing-kinds, and a fixed and known and named variety of causal properties, let us use this form as a means of calling attention both to certain distinctions among causal properties, and to certain additional questions of philosophical interest.

41. It is clear, to begin with, that the above formula imposes serious restrictions on the sort of thing that is to count as a ‘dispositional property.’ The first thing to note is that it captures only one of the uses of the terms we are considering, for we not only speak, as above, of an individual thing as having a certain causal property; we also ascribe causal properties to thing-kinds.

When we say of salt, for example, that it is (water-) soluble, we are

* Or, perhaps, ‘\( f(x, s, t) \)’—thus ‘\( x \) is red at place \( s \) and time \( t \).’

* For a careful analysis of dispositional concepts along these lines, see Burks (6).

The above account, however, was independently deve...
Wilfrid Sellars

clearly not ascribing this property to the thing-kind at a time. This
would, indeed, be logical nonsense. The statement that salt is soluble
has the form ‘D(K)’ and not ‘D(K,t)’. How, then, shall we represent
the connection between the two type levels? Shall we say
\[ D(K) \equiv (x) K(x) \implies D(x) \]
or
\[ D(K) \equiv (x)(t) K(x) \implies D(x,t) \]
Actually, of course, these two formulas are equivalent, provided we
stipulate that
\[ D(x) \equiv (t) D(x,t) . \]
But the question calls attention to the fact that not only do we say
“salt is soluble” rather than “salt is soluble at t,” we also say “This is
soluble” rather than “This is soluble at t,” or, to put it somewhat dif-
fently, “This is soluble” does not have the force of “This is soluble
now.” This suggests that the schema
\[ D(x,t) \equiv \Phi(x,t) \longrightarrow \psi(x,t) \]
does not do justice to the logic of words for causal properties in ordi-
mary usage. Just as something is of a kind, period, and not of a kind
at a time, so something has a causal property, period.* As a matter of
fact, this seems to be the heart of the distinction between properties
(causal or otherwise) and states. Expressions for states have the form
‘F(x,t)’; those for properties the form ‘F(x)’. And what would cor-
respond in ordinary usage to the ‘D(x,t)’ of the above schema is ‘(the
state of)’ being such that if it were \( \Phi \)-ed, it would \( \psi \).’ Thus, being magnetized is a state.† We implied above that not all properties are causal
properties. The point will be developed shortly. It must now be added,
though the point is less likely to be controversial, that not all states
have the form ‘(the state of)’ being such that if it were \( \Phi \)-ed, it would
\( \psi \).’ The state of being red at a certain time would seem to be a good
example of what might be called an “occurrent” as contrasted with a
“causal” state.

* For a related distinction, see Bergmann (3).
† Yet it would be an oversimplification, of course, to say that the form ‘F(x,t)’
represents exactly the force of ordinary expressions to the effect that a certain thing
is in a certain state. For when we say, for example, of a certain object that it is red,
or magnetized, we imply that it has been and will continue to be red or magnetized
for an unspecified period; otherwise we would say, “It is red (or magnetized) now”
or “... for the moment.”

counterfactuals, dispositions, causal modalities

42. Now if the above remarks are sound, they highlight anew the
central role in causal discourse of thing-kind concepts. For this is
soluble, period, rather than soluble-at-t, just because we are thinking of
this as belonging to a soluble thing-kind. But to make these remarks
stick, we must draw certain distinctions. For, to begin with, it might
be said that malleable is a causal property, and yet a thing can be
malleable at one time but not malleable at another. The rough and
ready answer to this objection is that the term ‘malleable’ is ambiguous,
and that in one sense of ‘malleable’, malleability may be a causal prop-
erty of iron, while in another sense it may be a state of this (piece of)
iron, which was not (in this sense) malleable a moment ago. But to
spell this out calls for some remarks on the notion of a capacity.

43. It is not my purpose to botanize causal characteristics, to draw,
for example, the familiar distinctions between ‘active’ and ‘passive’
powers, between quantitative (metrical or non-metrical) and non-quan-
titative causal properties, or between causal characteristics of various
levels (illustrated by the distinction between being magnetized and
being magnetizable). For the philosophical perplexities with which we
are concerned arise when the attempt is made to understand even the
most ‘elementary’ members of this family. Thus, our immediate pur-
poses will be achieved by reflection on the distinction between a dis-
position (as this term is currently used) and a capacity.

As a first approximation, this distinction can be put by saying that
to say of a certain kind of thing that it has the capacity to \( \psi \), \( \gamma \psi \), is to say that there is a combination of a circumstance, \( C \), and a something
done, \( \Phi \), which results in the \( \psi \)-ing of that kind of thing. Thus, roughly,
\[ \gamma \psi \equiv \mathfrak{D} (\exists C) (\exists \Phi) : (x)(t) : K(x) \cdot C(x,t) \longrightarrow \Phi(x,t) \supset \psi(x,t) \]

This reference to circumstances calls mind the fact that on our
account of dispositions to date, to ascribe a disposition to a thing is to
imply that the thing is actually in a favorable circumstance, the C of
‘(x)(t) K(x) \cdot C(x,t) \longrightarrow \Phi(x,t) \supset \psi(x,t)’. It might not have been
in C (or in any other favorable circumstance, \( C' \)), in which case a pre-
supposition of the ascription would not obtain. This means, however,
that we must reconsider the schema
\[ D(K) \equiv (x)(t) K(x) \implies D(x,t) \]
which tells us that if K has D, then any thing of kind K is always in a
circumstance such that if it were $\Phi$-ed, it would $\psi$. This suggests that what we want, instead, is the function $D_c(K)$, where
$$D_c(K) \equiv (x)(t) \ K(x) \cdot C(x,t) \cdot \text{implies } D(x,t)$$
which amounts to the concept of being disposed to $\psi$ if $\Phi$-ed when the circumstances are $C$. Only if things of kind $K$ would $\psi$ if $\Phi$-ed in any circumstances whatever, would it be proper to ascribe to $K$ a property of the form $D(K)$ as contrasted with $D_c(K)$.

44. The next point I wish to make can best be introduced by considering an example which takes us somewhat away from the 'ideal' universe of discourse in which we have been operating. It takes its point of departure from the fact that according to our original account of disposition terms, to ascribe a disposition to an object at a certain time, thus, "$D(x, t_1)$", is to imply the existence of a general fact such that if one knew it, and if one knew a certain fact about the circumstances of $x_1$ at $t_1$, and if one knew that $x_1$ had been $\Phi$-ed at $t_1$, one would be in a position to reason

$$x_1 \text{ was a } K$$
$$x_1 \text{ in } C \text{ at } t_1$$
$$x_1 \text{ was } \Phi-\text{ed at } t_1$$
$$\text{So, } x_1 \text{ } \psi-\text{ed at } t_1$$

And this raises the question, Are we not often in a position to ascribe with reason a disposition to an object, although we are not in a position to subsume the object under a generalization of the appropriate form—that is to say of the form, 'Ks $\psi$ when $\Phi$-ed in $C$?'

Suppose, for example, that the only chemical substances we have so far found to dissolve in water are salt, sugar and a few others, and that they all share the property of forming large white crystals. And suppose that by a chemical reaction which we can repeat at will, we produce large white crystals which consist neither of salt, sugar, nor any of the other substances in our list of soluble chemicals. Suppose, finally, that in the absence of any reason to the contrary we conclude that this new substance (do we also conclude that it is a substance?) is soluble.

Now, one way of interpreting the above reasoning is by saying that from the idea that all known soluble substances share the property of forming large white crystals, we drew the inductive conclusion that the product of this chemical reaction, as also having the property of forming large white crystals, (probably) belongs in its turn to a soluble

**COUNTERFACTUALS, DISPOSITIONS, CAUSAL MODALITIES**

thing-kind, though we did not yet know what soluble thing-kind (what substance). We were not, the interpretation would continue, in a position to reason,

```
This is a K
So, if it were put in water, it would dissolve
```

though we were in a position to reason,

```
This has the property $\lambda$
Things having the property $\lambda$ (probably) belong to a soluble thing-kind
So, if this were put in water, it would (probably) dissolve.
```
(where $\lambda$ is the property of forming large white crystals).

'But,' it might be asked, 'instead of concluding that if something has the property $\lambda$, then it (probably) belongs to a soluble thing-kind, i.e. that

$$(x) \lambda(x) \text{ implies (in all probability) } (\exists K) \ \text{Sol}(K) \cdot K(x)$$

why not conclude that $\lambda$ itself is a soluble thing-kind? The suggestion is, in other words, that the proper conclusion of the above inductive inference, instead of being that having the property $\lambda$ implies that there is a thing-kind $K$, such that

$$(x)(t) \ K(x) \cdot \Phi(x,t) \supset \psi(x,t)$$

should rather be, simply, that $\lambda$ is itself the thing-kind in question, a thing-kind of which salt, sugar, etc., are sub-kinds. Our thing-kind generalization would be

$$(x)(t) \lambda(x) \cdot \Phi(x,t) \supset \psi(x,t)$$

and 'Sol($\lambda$)' would be the counterpart of 'Sol(salt)'.

45. The answer is, in principle, straightforward. Words for thing-kinds not only embody a great deal of empirical knowledge, which is obvious, but they have quite a different role in discourse from that of expressions for properties, e.g., 'forming large crystals', or 'being white'. We might be tempted to locate this difference between, say, 'salt' and 'x' by saying that while they both stand for properties, the property of being saline fits into a scheme of classification which organizes our chemical knowledge into a perspicuous whole. And, of course, it is true that thing-kind words do go along with ways of classifying things. But this is only part of the story, and is quite misleading if taken for the whole.

To bring this out, let us suppose someone to ask, "Can we not
imagine that \( \lambda \) might have been the identifying property of a thing-kind; and if so, would we not then be in a position to say ‘\( x_1 \) is (a sample of) \( \lambda \)’ as we now say ‘\( x_1 \) is (a sample of) salt?’ But the change from ‘\( x \) is \( \lambda \)’ (‘\( x \) is large crystal forming’) to ‘\( x \) is a \( \lambda \)’ gives the show away. For if ‘\( \lambda \)’ were to acquire this use, it would no longer be the same term, as when John the miller became John Miller, the word “miller” acquired a new use.*

46. Words for thing-kinds have a special logic which is ill-represented by the schema

\[
K(x) = \nu t \left( P_1(x,t) \cdot P_2(x,t) \cdots P_n(x,t) \right)
\]

(where ‘\( P \)’ is a neutral term which ranges over both ‘causal’ and ‘non-causal’ characteristics.) There is, indeed, a sense in which a thing-kind word ‘means’ a certain set of characteristics; but, as elsewhere, the term ‘means’ is unilluminating. One must get down to cases and see just how words for thing-kinds are related to words for causal properties and to words for such items as color, shape, and number of legs.

Now, the basic flaw in the above schema is that it assimilates the logic of thing-kinds to that of complex properties. The point is not simply that thing-kind concepts are vague in a way which makes inapplicable the model of a set of separately necessary and jointly sufficient defining criteria. For while there are important differences between thing-kind and property expressions with respect to the applicability of this model, the difficulty of finding separately necessary and jointly sufficient criteria is not limited to expressions for thing-kinds. Indeed, the problem of vagueness had been discussed for some time in terms of such examples as ‘bald’, before the pervasiveness of the problem became apparent. The point is the more radical one that the relation of a thing-kind word to the criteria for belonging to that kind of thing is different in principle from the relation of words for characteristics of things to the criteria for the presence of these characteristics. “Lemon” and “bald” may both be vague, but they are so in radically different ways.†

* For a further discussion of thing-kind expressions which construe them as common names of the individuals belonging to the kinds, and compares the irreducibility of a thing-kind to its ‘criterion characteristics’ with the irreducibility of proper names to the definite descriptions which are the criteria for their application, see my “Form and Substance in Aristotle: an Exploration,” Journal of Philosophy, forthcoming, fall 1957.

† For basic discussions of vagueness and ‘open texture,’ see Kaplan (10), Kaplan and Schott (11), Pap (13), Waisman (20), and Wittgenstein (22).

47. One way of attempting to put this point is by saying that words for kinds of thing continue, in an important sense, to have the same meaning in spite of significant changes in their so-called defining traits, whereas words for characteristics do not. Thus, one might begin by claiming that a shift in the criteria for baldness would amount to the substitution of a new concept for the old one. Suppose, for example, that evolution diminishes man’s initial endowment of hair; might not the word “bald” come to be so used that one would have to have less hair than today in order to be called bald? And would not the word have changed its meaning? This, of course, is much too simple. Yet the very way in which it is too simple throws light on thing-kind words. For the term ‘bald’ is not a cold descriptive term; bald people are not merely people with a ‘small’ amount of hair, nor is it simply a matter, say, of having a small proportion of the original endowment (in which case there need have been no change of meaning to begin with). The logic of ‘bald’ involves the idea that being bald is not the sort of thing one would choose; and this theme is a continuing theme in its use.

48. This idea of a continuing theme illuminates the ‘meaning’ of thing-kind words. It is the role of these words in explanation which accounts for the fact that it can be reasonable to say “That wasn’t really gold” in spite of the fact that the object in question was correctly called gold according to the criteria used at the time the claim that is being disputed was made. And the statement that the object really wasn’t gold is not to be construed as a queer way of saying that the word “gold” no longer means exactly what it did. It is the regulative connection of thing-kind words with the schema

\[
K_\psi \text{ when } \Phi \text{-ed, in appropriate circumstances}
\]

which guides them through the vicissitudes of empirical knowledge.

This feature of the tie between a thing-kind word and the criteria by which one identifies members of the kind throws new light on the logic of general truths concerning thing-kinds. If thing-kind words were adequately represented by the schema

\[
K(x) = \nu t \left( P_1(x,t) \cdot P_2(x,t) \cdots P_n(x,t) \right)
\]

then general truths of the form

\[
(x) \ K(x) \cdot \text{implies} \cdot (t) \ \Phi(x,t) \supset \psi(x,t)
\]
would, viewed more penetratingly, have the form

\[(x)(t) \ P_1(x,t) \cdot P_2(x,t) \cdot \ldots \cdot P_n(x,t) \implies \Phi(x,t) \supset \psi(x,t)\]

Now, if \(P_1, P_2 \ldots P_n\) were ‘occurrent’ rather than ‘causal’ characteristics, there would be relatively little in all this to puzzle us. But it is clear on reflection that the criteria for belonging to a thing-kind are by no means limited to non-causal characteristics; and once we realize this, we begin to be puzzled.

For, we wonder, how is it to be reconciled with the idea, on which we earlier laid such stress, that when a soluble object dissolves in water, the fact that it is put in water causes it to dissolve? We took this at the time to mean that while there is no general causal fact to the effect that anything put in water dissolves, the soluble object has some character such that it is a general fact that anything having this character which is also put in water, dissolves. If we now ask, Can this character be another causal property or set of causal properties, or include a causal property? we are strongly tempted to say No—partly because we are tempted to think of the fact that the object has this additional character as a part cause (and hence of the fact that it is put in water as really only itself a part cause of the dissolving), and then to wonder how a causal property can be a cause; and partly because we smell the beginnings of a circle. If, on the other hand we try to fall back on the idea that the general fact in question is of the form

\[(x)(t) \ O(x,t) \implies \Phi(x,t) \supset \psi(x,t)\]

(where \(O\) neither is nor includes a causal characteristic), we are confronted by the brute fact that we don’t know any such general facts.

49. Now part of the solution to this puzzle consists in recognizing that even if the fact that the object was put in water is not the complete explanation of the fact that it dissolved, the putting in water was not, for this reason, only a part cause of the dissolving. Thus, when we explain the fact that a piece of salt dissolved in water by calling attention to the fact that it was a piece of salt, we are not implying that being a piece of salt is a part cause (along with being put in water) of the dissolving. More, indeed, must be known of an object than the mere fact that it was put in water, in order to infer that it dissolved. But such thing-kind generalizations as “Salt dissolves in water” include this more not by specifying additional part causes, but by restricting their scope to identifiable kinds of thing, in identifiable kinds of circumstance.

50. But the philosophically more exciting part of the solution consists in distinguishing between the causal properties of a certain kind of thing, and the theoretical explanation of the fact that it has these causal properties. For while causal generalizations about thing-kinds provide perfectly sound explanations, in spite of the fact that thing-kinds are not part-causes, it is no accident that philosophers have been tempted to think that such a phenomenon as salt dissolving in water must “at bottom” or “in principle” be a “lawfully evolving process” describable in purely episodic terms. Such an “ideal” description would no longer, in the ordinary sense, be in causal terms, nor the laws be causal laws; though philosophers have often muddied the waters by extending the application of the terms ‘cause’ and ‘causal’ in such wise that any law of nature (at least any nonstatistical law of nature) is a ‘causal’ law.*

It would be a serious mistake to think that a mode of explanation, in particular, ordinary causal explanation, which enables us to give satisfactory answers to one family of questions, cannot be such as by its very nature to lead us on to new horizons, to new questions calling for new answers of a different kind. The plausibility of the ‘positivist’ interpretation of theoretical entities rests on a failure to appreciate the way in which thing-kind generalizations by bunching rather than explaining causal properties point beyond themselves to a more penetrating level of description and explanation; it rests, that is, on a failure to appreciate the promissory note dimension of thing-kind expressions. The customary picture of the relation of ‘observational’ to ‘theoretical’ discourse is upside down. The ‘primacy’ of molar objects and their observable properties is methodological rather than ontological. It is the ultimate task of theory to re-create the observational frame in theoretical terms; to make available in principle a part of itself (at a highly derived level, of course) for the observational role—both perceptual, and, by containing a micro-theory of psychological phenomena, introspective.†

51. We said above that the picture of the world in terms of molar things and their causal properties (a) points beyond itself to a picture of the world as pure episode, and (b) leads, by its own logic, to the

---

* Cf. Collingwood (8), Feigl (9), and Russell (23).
† For an elaboration of this theme, and a more complete discussion of the nature and status of theories, see my essay “Empiricism and the Philosophy of Mind,” in Volume I of Minnesota Studies in the Philosophy of Science, particularly sections 39–44 and 51–55.
introduction of unobserved entities. It is important to see that these
two ‘demands,’ though related, do not coincide. For micro-theories
themselves characteristically postulate micro-thing-kinds which have
fundamentally the same logic as the molar thing kinds we have been
considering. And if they do take us on the way to a process picture of
the world, they do not take us all the way. For even if a ‘ground floor’
theory in terms of micro-micro-things were equivalent to a pure process
theory by virtue of raising no questions concerning the causal prop-
erties of these micro-micro-things to which it could not provide the
answer, it would not for that reason be a pure process theory. For the
logical form of a thing theory is, after all, characteristically different
from that of a theory whose basic entities are spatio-temporally related
events, or overlapping episodes.

52. The conception of the world as pure process, which is as old as
Plato, and as new as Minkowski, remains a regulative ideal; not simply
because we cannot hope to know the manifold content of the world
in all its particularity, but because science has not yet achieved the
very concepts in terms of which such a picture might be formulated.
Only those philosophies (New Realism, Neo-Thomism, Positivism, cer-
tain contemporary philosophies of common sense and ordinary usage,
etc.) which suppose that the final story of “what there is” must be
built (after submitting them to a process of epistemological smelting
and refinement) from concepts pertaining to the perceptible features
of the everyday world, and which mistake the methodological depend-
ence of theoretical on observational discourse for an intrinsically second-
class status with respect to the problems of ontology, can suppose the
contrary.

53. Important though these broader implications of our analysis may be,
to follow them up would take us far beyond the scope of this
essay. Our task is to quarry some of the stones for this more ambitious
enterprise. Let us, therefore, conclude the present section of this essay
with some remarks on the distinction between the identifying traits and
the properties of thing-kinds.

Suppose, to begin with, that a certain kind of thing, \( K_t \), has a certain
causal characteristic, \( P_1 \), which is not one of its identifying traits. One
might be tempted to think that the assertion that \( K_t \) has \( P_1 \) is equiva-
 lent to the assertion that anything having the traits by which \( K_t \) is iden-
tified, has \( P_1 \). If our argument to date is sound, however, this idea would

be incorrect—as assimilating thing-kind expressions to expressions for
complex characteristics. On the other hand, it is quite true that the
momentary cash value of the idea that \( K_t \) has \( P_1 \) is the idea that the
traits by virtue of which \( K_t \) is identified are conjoined with \( P_1 \). It is
important to note, however, that not all the identifying traits of a thing-
kind need be directly relevant to its possession of a given causal charac-
teristic. Indeed, we are often in a position to formulate generalizations
which, prima facie, have the form:

\[
(x) \cdot P_n \cdot \ldots \cdot P_{n+m}(x) \quad \longrightarrow \quad (t) \cdot \Phi(x,t) \supset \psi(x,t)
\]

where \( P_n \ldots P_{n+m} \) constitute a proper subset of the identifying
traits of a thing-kind, and where this group of traits may be found in
several thing-kinds. The existence of generalizations having (prima
facie) this form encourages the mistaken idea that thing-kind general-
izations are special cases of such generalizations, cases in which the
set of traits in question has acquired the status of a thing-kind concept
by virtue of being given a place in a classificatory scheme.

The truth of the matter is rather that the generalizations represented
above exist within a framework of identifiable thing-kinds, and that far
from it being the case that generalizations pertaining to thing-kinds are
a special case of generalizations pertaining to sets of characteristics,
the latter are abstractions from generalizations pertaining to thing-kinds,
and are, implicitly, of the form:

\[
(x)(K) : \Rightarrow K(x) \text{ implies } P_n \ldots P_{n+m}(x) : \text{ implies :}
\quad (t) \cdot \Phi(x,t) \supset \psi(x,t)
\]

54. Finally, what of the case where a causal property is one of the
identifying traits of a thing-kind? What are we to make of the formula:

\[
(x)(t) \cdot K(x) \cdot C(x,t) : \text{implies } \Phi(x,t) \supset \psi(x,t)
\]

(where \( \psi \)-ing when \( \Phi-ed \) (in \( C \)) is an identifying trait of \( K \))? We seem
to be confronted with a dilemma. Either the remaining identifying traits
of \( K \) imply the presence of this trait—in which case there is no point
to including it among the identifying traits—or they do not, in which
case how can a generalization of the above form be anything other than
a tautology? But surely when we suggested that

If \( x_i \) were \( \Phi-ed \), it would \( \psi \)

points beyond itself to

\[
F(x_i, \text{now}), \text{and } (x_i, \text{now})
\]