ON THE LOGIC AND CRITERIOLOGY OF CAUSALITY

Marc LEMAN

1. Introduction

In the present paper I argue in favour of a dynamical constructivistic counterfactual logic, which I propose to be the basis for our reasonings about causal relations.

In the first part of this paper I discuss a number of problems which are related to the classical conditional and counterfactual analysis of causation: (i) the interpretation of counterfactual logic, (ii) the status of laws, (iii) the direction of conditionship, (iv) the epistemological supplementation of the logical analysis. It is argued that the separation between the logical analysis on the one hand and the epistemological foundation on the other hand, leads to an ambiguous conception about the relation of cause and effect, and that the interpretation of causal counterfactuals involves problems with relation to the criteria for their evaluation.

In the second part of this paper, a dynamical conditional logic is proposed which inherently refers to the epistemological considerations concerning the causal relation. The evaluation of causal counterfactuals is proposed in terms of constructivism and, finally, some general remarks are made about the function of laws in the justification of causal connections.

2.The interpretation of Stalnaker-Lewis counterfactual logic

Counterfactual justification

Justification begins with a question: two events a and b occurred, how do we have to proceed in order to justify that there was a causal connection between these two events?

Following the familiar counterfactual approach, one could proceed as follows:

(i) Assume that a caused b; then it must be true that if a had not occurred, b would not have occurred.

Taking "A > B" as the counterfactual statement (if A had been the case, then B would have been the case) and "Ox" as the abbreviation of the statement "the occurrence of x", we may write:

(ii) a caused $b \supset . \sim Oa > \sim Ob$

On the assumption that a caused b, by showing that b would not have occurred if a had not occurred, i.e. by showing that Ob is counterfactually dependent on Oa, we have an argument in favour of the fact that a was the cause of b.(1) The next question is then: when do we say that the causal counterfactual "$\sim Oa > \sim Ob$" is true?

According to the counterfactual theory of Lewis, the statement "$\sim Oa > \sim Ob$" is true if and only if $\sim Ob$ is true in the nearest possible world(s) where $\sim Oa$ is true. Or in other words, in those possible world(s) where $\sim Oa$ is true and which resembles most to the world in which we suppose that a caused b, $\sim Ob$ must be true. Again we have to ask ourselves how we could proceed in order to show this. Two approaches seem sensible: The first goes back to the original interpretation of Stalnaker and Lewis, which is based on the notion of comparative similarity. The second is based on an interpretation in terms of complex regularities as the account of Mackie suggests.

Complex regularities

The interpretation of the counterfactual implication in terms of complex regularities is confined to the so-called *causal* counterfactuals, i.e., counterfactuals which express a causal dependence between two events.

According to Mackie (1974), the causal connection between two events is an instance of *some* probably complex regularity. It is assumed that some perhaps as yet quite unknown and unsuspected uniformity is instantiated. According to this interpretation, the infe-

(1) See STALNAKER (1968, LEWIS (1973a) for a theory of counterfactuals. See Lewis (1973b) and Swain (1978) for a counterfactual analysis of causation. I skip some details such as the requirement that a and b must be distinct, the possibility of having causal chains, problems concerning overdetermination and others. To the problem of the direction of conditioship I return.

rence from "~O*a* would be the case" to "~O*b* would be the case" is legitimated with reference to a complex regularity of the form O*a* & C ⇒ O*b* (if the occurrence of *a*, together with a set of true statements C would be the case, then the occurrence of *b* would always be the case).(2)

In other words: O*a* & C is a sufficient condition for O*b*, but not a necessary condition (there could be a D such that O*a* & D is sufficient for O*b* and D ⊬ C). By this, O*a* is a (though *I*nsufficient) *N*ecessary part of a (though *U*nnecessary) *S*ufficient condition for O*b*. O*a* is an *INUS*-condition which is also expressed by the statement "~O*a* > ~O*b*" (on the assumption that it is true).

Complex regularities of the afore-mentioned form warrant the truth of causal counterfactuals. They are, according to Mackie, the basis for our reasonings about possible worlds. For if we know that such a law is true in the actual world, we may deduce that in the possible world which resembles most the actual world, that *b* has not occurred if *a* has not occurred. Notice moreover that the laws of the actual world are the same in every possible world that is accessible from the actual world. This involves a reduction of the set of accessible worlds to those which have the same laws as the actual world. This principle is called the *principle of legal conservatism* (see e.g. Pollock (1981)).

The relation to Goodman

This way of reasoning about causal counterfactuals is very close to the well-known analysis of Goodman (1946). According to Goodman, A > B is true if and only if (i) there is a set of true statements C, such that A is co-tenable with C and there is a law A & C ⇒ B and (ii) there is no set of true statements D, such that A is cotenable with D and there is a law A & D ⇒ B.(3) In this approach, counterfactual inference is legitimated with reference to a law too.

The approach involves a circularity, for the definition of co-tenability is in terms of counterfactuals: A is cotenable with C = (def)

(2) See POLLOCK (1976), chap. III for an account of generalized subjunctive conditionals.

(3) LOEWER (1979) showed that Goodman's counterfactual logic is, apart from a few modifications, a subsystem of Lewis counterfactual logic in terms of possible worlds. He also notes the difference in interpretations of counterfactual conditionals. Goodman refers to laws, Lewis to comparative similarity.

$\sim(A > \sim C)$. Therefore, reference is made to epistemology and the ultimate foundation of laws and counterfactuals relies on good inductive evidence. Mackie (1974, pp. 55-56) holds the same view in claiming that the justification of causal counterfactuals rests in fact on the use of general propositions

> "... which we take to be confirmed by observations of the actual world, but which we feel justified in extending beyond the instances in which they are confirmed not only to other actual instances but to merely possible ones which are related to the confirming instances in the same way that other actual instances would be."

Comparative similarity

In confining ourselves to the evaluation of causal counterfactuals, another approach could be vindicated which goes back to Lewis' original interpretation in terms of comparative similarities between worlds. This approach has the advantage of taking causal counterfactuals prior to complex regularities. In fact we need no complex regularities in order to evaluate causal counterfactuals. Mackie calls such a method "primitive" in contradistinction with the sophisticated method just mentioned, because it does not rely on the combination of inductive and deductive reasoning, but on imagination and analogy:

> "I have observed another situation very like the present one, in which (unlike the present one) no event of the X type occurred. I borrow features from that other situation to fill out my imaginative picture of the possible situation that in the present circumstances X did not occur. In particular, if no event of the Y type occurred in that other situation, I see the possible situation that X did not occur in the present circumstances as continuing without the occurrence of Y." (1974, p. 56).

However, the comparison of similarities in different situations is a difficult enterprise. The notion of comparative similarity involves a number of problems such as circularity (the comparison of worlds would involve in his turn counterfactuals), vagueness in the equilibration of dissimilarities and problems with the symmetry of causal

counterfactuals (Cfr. e.g. Fine (1975), Bowie (1979), Stern (1981)).

All these considerations seems to suggest that the interpretation of causal counterfactuals in terms of comparative similarity cannot be maintained. The notion of comparative similarity involves too many a problem for the evaluation of causal counterfactuals.

Furthermore, in his argumentation in favour of a sophisticated method, Mackie (1974, p. 79) states that "in the unsophisticated method, one could say that to be prone to make such imaginative moves is somewhat like having an unconscious belief that there is some underlying regularity in the world".

3. Some problems with the logical analysis

Having modelled the logical part of the two classical counterfactual and conditional analysis, we now turn to an attack against these models. I will show that the separation between the logical aspect and the epistemological aspect of causal justification, which is inherent to both views, leads to an ambiguous conception of causation. This becomes clear especially with respect to the problem of the direction of conditionship. First of all, however, let us turn to a problem of Mackie's account which has lead us to the repudiation of his interpretation in terms of complex regularities, viz., the status of these complex regularities.

The status of laws

Mackie's approach is very subtle on this point. On the one hand, he defends the priority of singular causal statements to complex regularities on the basis of epistemological considerations, but, on the other hand, he seems to suppose causal regularities in the world, in order to legitimate counterfactual reasoning.

He says that singular causal statements do not imply that the sequences about which they are made of are instances of regularities of any sort (p. 77). But with regard to causation "as it is in the objects", he seems to say that there are complex regularities in the world, although he needs to introduce a supplementary epistemological criterium in order to make a distinction between accidental

regularities and causal regularities. The supplementary criterium is causal priority.

Mackie's subtility consists in the fact that he argues in favour of a sophisticated method of justification instead of a primitive method of justification. But in arguing for the sophisticated method, he has to suppose the existence of complex causal regularities in order to sustain causal counterfactuals. At the same time he seems to reject a regularity theory and argues in favour of the priority of singular causal statements. He says that the regularity theory, even in its improved form (as he gives), is not a complete account of causation in the objects. The regularity-theory misses a distinguishing feature:

> "Some causal mechanism or continuity of process *may be* an additional and distinguishing feature of sequences in which the earlier item causes the later..." (1974, p. 86).

This distinguishing feature is the causal priority which, however, is an epistemological feature; and therefore singular.

Moreover, even in accepting the regularity theory, supplemented with the criterium of causal priority, complex regularities can never be known in completeness. They are necessarily elliptic, we never can find all the conditions that are sufficient for the effect. This is an important remark with respect to the justification of causal connections. Taking the example of Scriven (1975): Assume that you may know that *d*ropping your watch caused it to *st*op running. Do we know a complex regularity of the form "$Od \,\&\, C \Rightarrow Ost$" which would show the occurrence of this event together with a set of conditions to be sufficient for the fact that it stops running? Moreover, following Mackie's approach, you can never specify what elements of C are "essential". It is impossible to give a reliable exhaustive description of all the components in C, but you have to suppose that there exist such a law, even though one can't produce it. With Scirven I would call such complex regularities: phantom-laws.

I cannot agree with such a conception of laws, the origin of which ensues from a separation between the logic of reasoning and the epistemological foundation. The inferences on the basis of such laws seems to bear witness to a naïve conception of justification.

The direction of conditionship

The so-called problem of the direction of conditionship concerns the asymmetry of the causal relation and the symmetry of causal counterfactuals. Counterfactual logic cannot give an account of the asymmetry of causation on purely logical grounds.

Assume for example that the two events a and b occurred: a = John throws a brick through the window, b = the window breaks. Two statements may be formulated:

(i) If John would not have throwed the brick, the window would not have broken. ($\sim Oa > \sim Ob$)
(ii) If the window would not have broken, then John would not have throwed the brick. ($\sim Ob > \sim Oa$)

With our knowledge of breaking windows, it is easy to see that both (i) and (ii) are true in the Stalnaker-Lewis interpretation. And especially (ii): in the closest worlds where the window did not break, it is true that John did not throw the brick.

However, if "a caused b" is true, then "$\sim Oa > \sim Ob$" must be true, but in accepting that "$\sim Oa > \sim Ob$" is true, we have to accept that "$\sim Ob > \sim Oa$" is true, which would be an argument in favour of the fact that b caused a, contrary to our hypothesis. We therefore have to conclude that counterfactual dependency is not the only distinguishing feature of causation. How could we refine our analysis in order to grasp the asymmetry? Two proposals have been put forward, one which corresponds with the comparative similarity approach, the other fits in with the principle of legal conservatism.

(a) Lewis miracle analysis

Lewis (1973b and 1979) argues against the principle of legal conservatism and in favour of a refinement of the notion of comparative similarity. The counterfactual in (ii) is then conceived as a back-tracking counterfactual (it says that if the present were different, then the past causes would have to be different, else they would have caused the present to be as it actually is) to which the so-called standard conditions for the notion of comparative similarity must be applied. By this, back-tracking counterfactuals are eliminated. The basic principle of this method (which admits miracles) is grounded in

the idea that the past would *not* be different if the present would be different, the past causes would fail somehow to cause the same present effects.

Several objections may be raised against this approach. First of all, Lewis' proposal rests on an arbitrary proxy, for on what grounds do we conclude that (ii) is a backtracking conditional and (i) is not? Second, and in addition to this, this proposal concerning the direction of conditionals remains merely ad hoc, it does not even fit in with a *criterium* for the asymmetry. And this too is a feature we are looking for.

(b) Swain's alternative

Swain (1978) differs from Lewis in admitting causal dependencies in both directions ($\sim Oa > \sim Ob$ and $\sim Ob > \sim Oa$), but he adds a new condition that establishes the asymmetry. His proposal looks attractive for an approach such as Mackie's, because of the fact that the principle of legal conservatism may be maintained. However, several authors have critisized this proposal (see e.g. Davis (1980), Bunzl (1980)).

Perhaps the following proposal in terms of numeric dissimilarities would do the job. The idea behind it is based on the counting of dissimilarities.

The definition of cause would be: if a and b occurred, the a caused b if and only if:

(i) $\sim Oa > \sim Ob$

(ii) the world in which a occurred without b counts less dissimilarities to the actual world than the world in which b occurred without a.

However, the counting of dissimilarities seems to be a suspicious method, and even if this proposal could be held, the condition in (ii) would remain purely ad hoc.

Lewis' miracle analysis and Swain's alternative appear as a solution within the logical analysis. What is needed, however, is an epistemological foundation of the asymmetry.

4. The supplementation of the logical analysis: epistemological foundation

The need of an epistemological supplementation to the logical analysis is suggested by two considerations:

(a) the circularity of the regularity analysis (cfr. Mackie, Goodman) suggest a foundation on inductive grounds for the ultimate foundation of laws and counterfactuals relies on good inductive evidence.

(b) but even in repudiating the regularity approach (as a result of the considerations concerning the status of comples regularities) and arguing in favour of a primitive interpretation of the counterfactual connective, the problem of the direction of causal counterfactuals remains and requires a criteriological supplementation.

In this section two proposals are discussed, first the manipulative criterium of von Wright, second, the causal priority of Mackie.

von Wright's criterium in terms of manipulations

von Wright (1971, p. 70) supplements the conditional analysis with the following epistemological criterium:

> "I now propose the following way of distinguishing between cause and effect by means of the notion of action: p is a cause relative to q, and q an effect relative to p, if and only if by doing p, we could bring about q or by suppressing p we could remove q or prevent it from happening."

And in von Wright (1973, p. 107):

> "what makes p a cause-factor relative to the effect-factor q is, I shall maintain, the fact that by *manipulating* p, i.e., by producing changes in it "at will" as we say, we could bring about changes in q".

Manipulating events in the world is the basis for our knowledge of the causal relation. Therefore, von Wright's manipulative criterium could be the distinguishing feature we are looking for. The definition of cause world run as follows:

a caused b if and only if
(i) $Oa \,\&\, Ob$
(ii) $\sim Oa > \sim Ob$
(iii) by manipulating Oa, we could bring about Ob (or remove, or prevent it from happening).

Condition (iii) remains however a bit vague and has to be adjusted to our past tense of a caused b. We have to refer to a situation which is in all respects like the one in which a caused b, but in which a is not the case an b is not the case. Then, by manipulating Oa, we could bring about Ob.
Condition (iii) may then be reformulated as follows:

(iii*) $\sim Oa \,\&\, \sim Ob >$ (manipulating Oa would bring about Ob)

Though this approach seems attractive. I cannot accept it because of the separation between the logical analysis and the epistemological foundation. This separation renders von Wright's approach heterogeneous: On the one hand he analyses causal connections in terms of necessary and sufficient conditions that are supported by regularities. In this sense, his approach is Neo-Humean and agrees to Mackie's conditional analysis. On the other hand, the notion of a nomic connection is introduced on the basis of the epistemological criterium. The nomic connection is known through manipulative interferences in the world and has to do with the way causal relations are distinguished from accidental regularities. This seems to fit in with a dynamic view of causality in the sense of Bunge (1951) and Apostel (1974). The causal connection is there conceived as a production-relation in the world. I hesitate however to ascribe to von Wright such an approach for the simple reason that he takes the dependence of the causal relation upon the concept of human action as merely epistemological (not ontological).

Both the logical analysis and the epistemological foundation remain separated and give rise to an ambigues conception of causation. In our alternative proposal, we formulate a logical analysis which inherently refers to epistemology. In this respect it is interesting to note that condition (iii*) involves a counterfactual statement that is related to action: "doing something would bring about something". Could we give an account of such statements in a stringent logical form and base

our analysis of causal statements on it? Moreover, do we still need statement (ii) if we could give a clear account of (iii*)? In fact, what we want to say is that Ob would not be *produced*, if a had not occurred.

Mackie's causal priority

Mackie suggests that epistemological considerations involve a (certain) dynamic conception of causation. In order to give an account of this feature, he introduces the notion of causal priority in the logical analysis. As a result, the distinction between accidental regularities and causal regularities seems to be established, the problem of the direction of conditionship seems to be solved, and attention is paid for contiguity. The distinguishing feature of causal sequence is, then, the conjunction of necessity-in-the-circumstances with causal priority. This leads to the following formulation in terms of counterfactuals:

a caused b if and only if: $(Oa \,\&\, Ob)\,\&$
$(\sim Oa > \sim Ob)\,\&$
(Oa was causal prior to Ob)

According to Mackie, the notion of necessity-in-the-circumstances-plus-causal-priority involves the thought of what would happen when a possible world, constructed by some appropriate alteration from the actual world, was allowed to run on. It is stated that X is necessary in the circumstances for and causaly prior to Y provided that if X is put into the world in the circumstances referred to and the world runs on from there, Y will occur.

The notion of causal priority is based on two assumptions: (a) that merely possible worlds have an independent running, i.e., that worlds have some way of running from one change to another, and (b) the possibility that agents interfere in the world.

The assumption of the independent running of possible worlds presupposes that the actual world has some laws of working which can be carried over to the possible worlds. This fits in with the principle of legal conservatism.

The assumption that agents interfere in the world constitutes in fact the epistemological foundation of the notion of causal priority. As I already said with respect to von Wright's proposal, this involves another kind of dynamics which leans against a conception of causa-

tion in terms of production. Notice that the first assumption also involves a kind of dynamic, but this cannot be interpreted in terms of production.

It is interesting to see how Mackie tries to escape the action-theoretic foundation of causal priority by carrying on the analysis to a further stage and by reducing its dependence on the notion of agency to the notion of fixation.

As a result of this reduction, the causal priority means that it is not the case that O*b* is *fixed* before O*a* (on the assumption that *a* caused *b*). This means that it might be the case that O*a* is fixed before O*b* and also that it might be the case that *b* occurs from the moment on that O*a* is fixed. The difference between the action-theoretic foundation and the interpretation in terms of fixation is that the independent running of possible worlds allows for the introduction of events without refering to agents and without refering to a production-relation or a necessary connection.

But how would such a model look like? The first assumption gives a hint in a certain direction. Indeed, possible worlds may be conceived as world states in a history. The development of a theory of conditionals with tense is therefore very welcome: it would become possible to give an account for the direction of conditionship in a dynamical model.

The independent running of possible worlds would be conceived of as the ordening of world states into a treelike structure by a relation "earlier than": $<$; (see Thomason and Gupta (1980)). An equivalence relation "$\equiv$" picks out moment pairs that are to count as co-present. In this model, the Stalnaker selection-function, relativised to both a moment i and a possible history h, picks out a pair (i′, h′) in which the antecedent is true and which is closest to (i, h). The moment i′ must be co-present with i and the history h′ must pass through it. According to this model, the conditional $A > B$ is true at (i, h) if and only if B is true in the world (i′, h′) in which A is true. This model allows for the introduction of tense notions, such as F ("in the future it will be the case that") and P ("in the past it was the case that"). The definition of cause could now be reformulated as follows:

$$a \text{ caused } b \text{ if and only if } (Oa \,\&\, Ob)\,\& \; (\sim Oa \,\&\, \sim Ob > . \; (Oa > Ob) \vee (Oa > FOb))$$

This may be represented as in fig. 1.

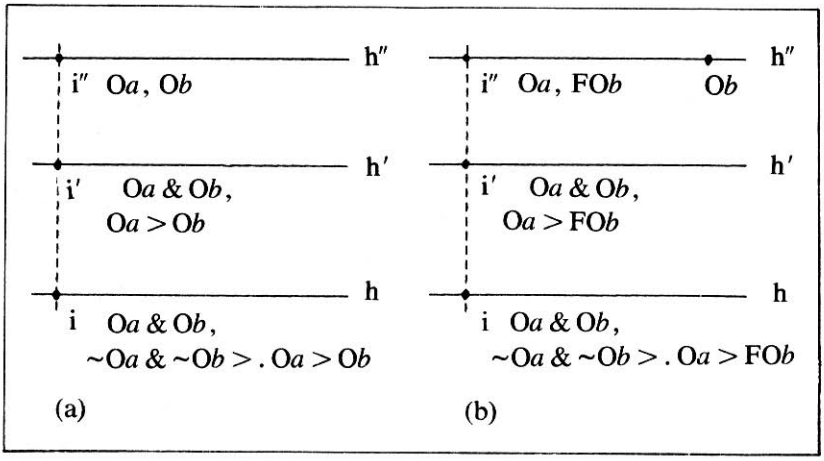

fig. 1

The truth of Oa in i″ could then be interpreted as the fixation of Oa, the bringing about of Oa into the world. From there on, as it is the case in (b), our world runs to the moment in which Ob becomes true.

This approach looks attractive for the Humean account: it is compatibel with Humean causation and seems to involve an aspect of dynamics. However, there are a number of problems. First, the state of the actual world has changed and the picking out of the closest world where $\sim Oa \,\&\, \sim Ob$ is true at moment i (i.e. "now") does not ensure (on the assumption that *a* caused *b in the past*) that "$(Oa > Ob) \vee (Oa > FOb)$ is true. The conditions that were true at the time when *a* caused *b* may have changed. This could be solved by ascribing the complex conditional to the past:

$$P\,(\sim Oa \,\&\, \sim Ob > .\, (Oa > Ob) \vee (Oa > FOb))$$

The second problem, however, seems more pervasive. The analysis in terms of temporal conditionals has the advantage of introducing a certain kind of dynamism into the logic of cause and effect. But the introduction of time as the distinguishing feature of causal asymmetry is not accepted by the greater part of the philosophers. I shall not go deeper into this discussion and shall confine myself to the remark that in taking this model as the basis of our analysis, a reduction is carried

out from the manipulative criterium to a temporal criterium.

One might argue that the truth of O*a* in i″ could be interpreted as the fixation of O*a* in i″, or even as the bringing of O*a* into the world i″. However, this interpretation is merely ad hoc. Whatever the notion of fixation may mean, it is clear that time is introduced as the criterium for causal asymmetry. It is said that the effect may not occur before the cause.

In addition to this, it may be objected that the introduction of time in the model is not inherent to the causal sequence itself. To the contrary, time is imposed onto the system and singular causal sequences are instantiations of general laws through time. The kind of dynamics that is expressed by means of temporal notions in this extended counterfactual approach is not inherent to singular causal connections. This is contrary to the epistemological considerations which suggest dynamics in singular cases.

5. Dynamical constructivistic counterfactual logic and causal justification

The separation between the logical analysis on the one hand and the epistemological foundation on the other hand, leads to an ambiguous conception about the relation of cause and effect. We also saw that the interpretation of causal counterfactuals involves problems in connection with the criteria for their evaluation.

In this section, an alternative approach is proposed which relies on a dynamical constructivistic counterfactual logic.

Dynamics

The development of a dynamical constructivistic logic must be seen in the context of the justification of causal connections between events. When two events *a* and *b* occurred, we justify a causal connection between them by referring to counterfactual statements.

In the light of the previous discussion, we want the counterfactual statements to give an account of the *dynamics* of causation in the sense of Bunge (1959) and Apostel (1974). This dynamical conception is suggested by epistemological considerations, for our manipulating events in the world, i.e., the way in which causal knowledge is

acquired, seems to suggest, contrary to Hume, a *producing force* in the world. The idea of a producing force or necessary connection may also give a better understanding of the asymmetry and the contiguity of cause and effect.

Some readers may object against the transition from epistemological considerations to ontological statements. However, I do not claim that there *is* a producing force in the world, but that, all arguments considered, the existence of such a force is the only valuable hypothesis.

The justification accounts for this dynamical concept of causation with reference to dynamical counterfactual statements, e.g., when two events *a* and *b* occurred, we say that O*b* would not be *produced* if *a* had not occurred.

In general, a causal counterfactual has the form "if A would be the case, then B would be *produced*". We may however interpret the production of B as the result of a procedure, say γ, which is carried out by some mechanism in the world. γ (causal force) may be seen as a binary relation between two states: it starts in the state in which A is true and stops in the state in which B becomes true. Dynamics may be obtained by introducing the operator "$[\gamma]$" to the consequent of the causal counterfactual:

$$A > [\gamma] B$$

to be read as "if A would be true, then γ would always lead to B. (For an account of dynamical logic, see Harel (1979).)

Following the Stalnaker-Lewis evaluation of counterfactual statements, "$A > [\gamma] B$" is true if and only if, in the closest world in which A holds, it is true that γ always leads to B. This may be represented as in fig. 2.

However, we rejected this analysis of counterfactual statements because of the vagueness of the notion of comparative similarity. But we rejected Mackie's evaluation in terms of complex regularities as well! We therefore have to develop a new method for the evaluation of causal counterfactuals.

A constructivistic evaluation of causal counterfactuals

Instead of selecting the possible world(s) in which the counterfactual antecedent is true, we require that the possible world in which it is

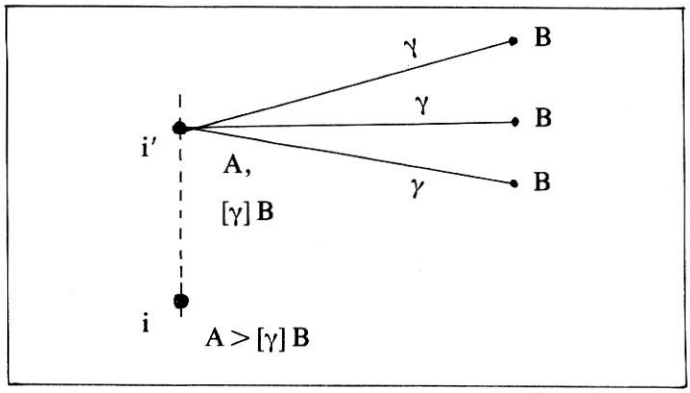

fig. 2

true must be constructible, i.e., that some finite procedure, say β, leads to that world. This procedure β may be conceived as a program, as a sequence of operations one has to carry out. The carrying out of the procedure may be called an *action*.

In order to specify the truth-conditions of a causal counterfactual, we have to require that the world which we obtain by carrying out the procedure β resembles in a sense the real (actual) world. We therefore have to make a description of the real (actual) world or at least of the conditions of the "partial history" which we consider relevant and which we want to keep constant in the new situation we want to construct.

It is clear that the constructivistic evaluation goes a step further than the method of comparing similarities between worlds, for it is required that some finite procedure must lead to the intended situation. Otherwise, when there is no procedure, the statement of the causal counterfactual is meaningless.

Notice that we do not claim that we reach the possible world (i.e. the world in which the antecedent is true), by making *minimal* changes to the real (actual) world. To the contrary, it is possible that we first have to change the real (actual) world in many respects before we are able to construct a situation in which the aforementioned conditions obtain together with A. Furthermore, in saying that the description of the real world depends on what we think to be relevant to the situation we want to construct, we are introducing pragmatical considerations.

It is convenient to consider the world in which the antecedent is true as *an experimental situation*: there is a set of relevant conditions and a set of possible outcomes.

We shall say that the causal counterfactual "$A > [\gamma] B$" is true in the actual situation, if and only if some procedure β leads to a situation in which A is true, which resembles the actual situation in the aforementioned sense, and which is such that γ always leads from that situation to a situation in which B is true.

In order to make clear that the connected evaluation is constructivistic, we shall write dynamical constructivistic causal counterfactuals as follows:

$$\beta A > [\gamma] B$$

This expression is to be read as: "if A would be the case, then B would be produced" and accepted to be true just in case it satisfies the aforementioned truth-conditions.

We now may define:

a is the cause of b if and only if
$\forall \beta_1 \exists \gamma_1 (\beta_1 Oa > [\gamma_1] Ob) \; \& \sim \forall \beta_2 \exists \gamma_2 (\beta_2 Ob > [\gamma_2] Oa)$

We shall define a to be the cause of b if and only if for every procedure β_1 there exist a force γ_1 such that if Oa would be the case, as a result of procedure β_1, then γ_1 would always produce Ob, whereas it is not the case that for every procedure β_2 there exist a force γ_2 such that, if Ob would be the case, as a result of a procedure β_2, then γ_2 would always produce Ob.

Causal justification

Thus far, we introduced a dynamical counterfactual logic and proposed a method for the evaluation of causal counterfactual statements. We now come to the question how this constructivistic evaluation enters into the justification of causal statements.

As I said already, when two events a and b occurred, we justify a causal connection between both events by stating that Ob would not be *produced* if a had not occurred. How do we have to represent and evaluate this counterfactual in terms of a dynamical constructivistic counterfactual logic?

Classical counterfactual analysis states that a caused b because "$\sim Oa > \sim Ob$" is true - together with an additional criterium that warrants the asymmetry -. By analogy, one might argue that a caused b because "$\sim Oa > \sim[\gamma] Ob$" is true: if a had not occurred, then γ would not always have lead to the occurrence of b. (Notice that "$[\gamma]$" warrants the asymmetry.) This, however, will not do, because it means in fact that if a had not occurred, then γ might have lead to a situation in which b does not occur: $\sim Oa > <\gamma> \sim Ob$.

Another proposal would be $\sim Oa > [\gamma] \sim Ob$: if a had not occurred, then γ would always lead to a situation in which b does not occur. This is, however, contrary to our intuïtions, for it means that there is some thing (a causal force) which *produces* the absence of events.

The way out is to say that "$\sim(\sim Oa > [\gamma] Ob)$" is true, i.e. that "$\sim Oa > [\gamma] Ob$" is false. However, the constructivistic evaluation may involve some troubles. First of all, it is clear that "$\sim Oa > [\gamma] Ob$" is not necessarily false in the *actual* situation, but it *is* necessarily false in the situation where a and b occurred. We have to say: "in the past, "$\sim Oa > [\gamma] Ob$" was false". Second, how do we have to evaluate this counterfactual? We could proceed by making a description of that past situation, and then seeking for a procedure which would lead, starting from the actual situation, to a new situation which resembles that past situation, but in which a does not occur. In that situation, we then may argue, the occurrence of b will never be produced. However, can we speak of an experimental situation when no events are manipulated?

In order to deal with such counterfactual statements, we may rely upon another kind of justification which is more sophisticated in the sense that referrence is made to the established knowledge. The sophisticated justification, which is intended to deal with the evaluation of "$\sim(\sim Oa > [\gamma] Ob)$", relies on "laws". By a law we here mean: the general resuls of our manipulating the world. They express the kind of general knowledge that is based on singular experiments. Laws have a dispositional nature which inherently refers to experimentation. Hence, in saying that "all P are $\hat{Q}$" is a law, it is meant that all things with property P have the disposition of being $\hat{Q}$. (See also Fetzer (1981).) Using Pollock's general conditional one may write:

$$P\mathrm{x} \Rightarrow \hat{Q}\mathrm{x}$$

To be read: for any x, if x would have the property P, then x would have the disposition of being $\hat{Q}$. Having the disposition of something is then defined as follows:

$$\hat{Q}\mathrm{x} =_{\mathrm{def}} \beta R\mathrm{x} > [\gamma]\, OQ\mathrm{x}$$

To be read: x has the disposition of being $\hat{Q}$ means, if a state R of x would be the case, then the outcome-response Q of x would be produced. Notice that the bringing about of the outcome-response Q of x is the reason why we say that x has the disposition of being $\hat{Q}$. For the evaluation of this counterfactual, we rely upon some procedure β.

The general characterisation of causal laws (in the aforementioned sense) is then:

$$P\mathrm{x} \Rightarrow (\beta R\mathrm{x} > [\gamma]\, OQ\mathrm{x})$$

To be read: for any x, if x would have the property P, then, if a state R of x would be the case (as the result of a procedure β), then the outcome-response Q of x would be produced. But there may be several ways of bringing about the same outcome-response. The full characterisation of a law is then stated as follows:

$P\mathrm{x} \Rightarrow \hat{Q}\mathrm{x}$ means: $P\mathrm{x} \Rightarrow .\ (\beta_1 R_1\mathrm{x} > [\gamma_1]\, OQ\mathrm{x}) \,\&$
$(\beta_2 R_2\mathrm{x} > [\gamma_2]\, OQ\mathrm{x}) \,\& \ldots \,\&$
$(\beta_n R_n\mathrm{x} > [\gamma_n]\, OQ\mathrm{x})$
where $R_1, R_2, \ldots, R_n$ refer to different states of x,
where $[\gamma_1], [\gamma_2], \ldots, [\gamma_n]$ refer to different producing forces and where $\beta_1, \beta_2, \ldots, \beta_n$ refer to the procedures for the evaluation.

We now dispose of the means to give a full account of causal justification which refers to laws. If both a and b occurred, we may say that a *was* the cause of b, if by reference to a *network* of laws, we have reasons to believe that b would not have been produced if a had not occurred.

To refer to a network of laws is to refer to (a set of) scientific theories. This involves the possibility of different levels of justifica-

tion. The appropriate level of a justification depends on a number of pragmatical considerations. I will not elaborate on this, but the idea behind it may be made clear with a simple example.

Assume that two events occurred: a = John strikes a match, b = the match lights. In order to show that both events are causally connected, and more especially that a caused b, we have to evaluate the counterfactual that if the match had not been struck, then the lightning of the match would not have been produced. We may proceed in two steps.

First we show that the striking of a match is a fricative act by which an object is rubbed against another object. Reference is made to a general law which states that objects have the disposition of becoming warm: objects are "warm-able":

$$(1)\ Jx \Rightarrow \hat{W}x$$

To be read: for any x, if x would be an object, then x would have the disposition of becoming warm. There are several ways of heating an object, but we are interested in one specific way, viz., by means of friction. We thus arrive at:

$$(2)\ \hat{W}x =_{(def)} \beta Fx > [\gamma] OWx$$

Whence the law has the form:

$$(3)\ Jx \Rightarrow (\beta Fx > [\gamma] OWx$$

To be read: for any x, if x would be an object, then, if x would be in a state of friction, then this would produce that x becomes warm.

Second, matches are made of a material that enflames at a certain temperature. This is expressed by the law:

$$(4)\ Sx \Rightarrow Ex$$

Again we have to fill in the specific testprocedure such that the material enflames, viz., we have to warm it up:

$$(5)\ Sx \Rightarrow (\alpha Tx > [\delta] OEx)$$

For any x, if x would be made of sulphur, then, if x would have a temperature of n° C, then the enflaming of x would be produced. The procedure α by which we arrive a the desired temperature refers to (1) and is stated in terms of friction. The outcome-response is the lightning of the match.

By stating such a network of laws with the help of the contextual information, we now may conclude that if the match had not been struck, then the lightning of the match would not have been produced.

Summarizing, we proposed a dynamical counterfactual logic as the basis for a logic of the causal relation. We provided a constructive method for the evaluation of these counterfactuals. However, in the context of causal justification, statements of the form "if *a* had not occurred, then *Ob* would not have been produced", may cause some trouble with respect to their evaluation. In order to deal with these counterfactuals, we may rely on a sophisticated procedure which refers to laws. The reliance upon dispositional laws and the implicit reference to scientific theories rules out a too naïve account of justification. (4)

Seminarie voor Logica en
Wijsbegeerte van de Wetenschappen Marc LEMAN
Rogier 44
9000 Gent

(4) I wish to thank Diderik Batens for reading and criticizing the first draft of this paper.

REFERENCES

Apostel L. (1974), *Matière et Forme*, Communication and Cognition, Gent.

Bowie G.L. (1979), "The similarity Approach to Counterfactuals: Some problems", *Noûs*, 13, pp. 477-498.

Bunge M. (1959), *Causality and Modern Science*, Dover Publications, New York.

Bunzl M. (1980), "Causal Preemption and Counterfactuals", *Philosophical Studies*, 37, pp. 115-124.

Davis W. (1980), "Swain's Counterfactual Analysis of Causation", *Philosophical Studies*, 38, pp. 169-176.

Fetzer J. (1981), *Scientific Knowledge*, D. Reidel Publ. Comp. Dordrecht.

Fine K. (1975), "Critical Notice of David Lewis, *Counterfactuals*", *Mind*, 84, pp. 451-458.

Goodman N. (1946), "The problem of counterfactual conditionals", in *Fact, Fiction and Forecast*, Havard University Press 1983.

Harel D. (1979), *First Order Dynamical Logic*, Springer, Heidelberg.

Leman M. (1983), *Causaliteit en Irrealissen*, University of Gent, mimeography.

Lewis D. (1973a), *Counterfactuals*, Basil Blackwell, Oxford.

(1973b), "Causation" in E. Sosa (ed.) (1975).

(1979), "Counterfactual Dependence and Time's Arrow", *Noûs*, 13, pp. 455-476.

Loewer B. (1979), "Cotenability and Counterfactual Logics", *Journal of Philosophical Logic*, 8, pp. 99-115.

Mackie J. (1974), *The Cement of the Universe*, Oxford at the Clarendon Press, 1980.

Pollock J. (1976), *Subjunctive Reasoning*, D. Reidel Publ. Comp. Dordrecht, Holland.

(1981), "A Refined Theory of Counterfactuals", *Journal of Philosophical Logic*, 10, pp. 239-266.

Scriven M. (1975), "Causation as Explanation", *Noûs*, 9, pp. 3-16.

Sosa E. (ed.) (1975), *Causation and Conditionals*, Oxford Univ. Press, Oxford.

Stalnaker R. (1968), "A Theory of Conditionals", in E. Sosa (ed.) (1975).

Stern C. (1981), "Lewis' Counterfactual Analysis of Causation", *Synthese*, 48, pp. 333-345.

Swain M. (1978), "A Counterfactual Analysis of Event Causation", *Philosophical Studies*, 34, pp. 1-19.

Thomason R. and Gupta A. (1980), "A Theory of Conditionals in the Context of Branching Time", in W. Harper et all. (eds.), *Ifs*, pp. 299-322, D. Reidel Publ. Comp., 1981.

von Wright G.H. (1971), *Explanation and Understanding*, Routledge & Kegan Paul, London.

(1973), "On the Logic and Epistemology of the Causal Relation", in E. Sosa (ed.) (1975).