

## Do effects counterfactually depend upon their causes?

---

### Introduction

#### 1 The Counterfactual Analysis of Causation

##### 1.1 The problem of counterfactuals

##### 1.2 The theory of Possible World Semantics

##### 1.3 Conclusion

#### 2 Actual Objections

##### 2.1 Early problems

###### 2.1.1 Epiphenomena

###### 2.1.2 Pre-emption

###### 2.1.3 Trumping pre-emption

##### 2.2 Fine and the counterintuitive falsity of counterfactuals

##### 2.3 Menzies and probabilistic theories of causation

#### 3 Conclusions

Giovanni C. Lorusso  
gioloru@yahoo.it

## **Do effects counterfactually depend upon their causes?**

*“I think it is time to give up and try something else”*

Theories of causation have always been built on *regularity* analysis. However, as Lewis pointed out, it “tends to confuse causation itself with various other causal relations”.

Indeed for Lewis there is a need of converging (by making them similar) the ontology of actual and possible events.

This becomes possible following the Theory of Counterfactuals as explained in his paper (*Lewis, 1973*). However several difficulties arise. Therefore, a possible synopsis is the following: first I will briefly define the Counterfactual Theory (1); I will then summarize some actual objections concentrating on Lewis’ latest answers (2) and showing that by remarking the analysis of *sensitivity* and *transitivity* (3) most of those objections (in particular related to the possibility of intrinsic relations between events) can be properly avoided and the possibility of an *asymmetric counterfactual dependence* between cause and effect is still possible.

### **1 \_ The Counterfactual Analysis of Causation**

#### **1.1 \_ The problem of counterfactuals**

The reason counterfactual statements have always stirred up skepticism within empiricist philosophers lays on the fact that those conditionals appear to refer to *non-actual possibilities*. For this reason earlier theories of counterfactuals were strictly *meta-linguistic* theories.

However, if we can define the similarity within these possibilities and the actuality of what really happened, we are then able to state the ontological value that appear to be causal dependent between cause and effect in order to explain the fact that: “we think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it”<sup>1</sup>.

#### **1.2 \_ The theory of Possible World Semantics**

Lewis remarks that in order to comprehend counterfactuals we must define them as statements about possible alternatives to the actual

---

<sup>1</sup> Lewis 1973, p161

situation. Therefore, the first step we make is by considering the *primitive relation* of similarity among *possible worlds*<sup>2</sup>.

By the application of this relation we can define that:

A  $\square \rightarrow$  C is true at a world  $w$  iff either (i) there are no possible A-worlds (vacuous truth) or (ii) some A-world where C holds is closer to  $w$  than is an A-world where C does not hold<sup>3</sup>.

So that, we don't assume that there must be at least one closest A-world, but if there is any, then (ii) holds.

We can now define causation in terms of counterfactual dependence by determining the truth condition in the following steps:

- A) Defining the notion of *counterfactual dependence*;
- B) Defining the notion of *causal dependence* by the usage of counterfactual dependence;
- C) Defining *causation* by the usage of causal dependence.

In (A) we determine counterfactual dependence between two classes of possible alternative propositions where every element of the first class can determine the antecedent of a counterfactual (that has to be true) in which its consequent is an element of the second class<sup>4</sup>.

But while by counterfactual dependence we determine a relation between sets of propositions, in (B) we define a relation between events. Indeed causal dependence is an extension of the notion of counterfactual dependence applied to events. This is possible by considering that for every event there is a corresponding proposition so that for every  $e$  there is a corresponding proposition  $O(e)$ . Then,  $O(e)$  will be true in every possible world in which  $e$  happens and false in every possible world where  $e$  doesn't happen.

We can then conclude that a relation of *causal dependence* between distinct events  $a$  and  $b$  is a relation of counterfactual dependence between  $O(a)$  and  $O(b)$ <sup>5</sup>. In addition, in order to simplify the impossibility of compossible propositions, we can define that  $e$  causally depends on  $c$  iff the set  $\{O(e), \sim O(e)\}$  counterfactually depends on the set  $\{O(c), \sim O(c)\}$ .

Doing so, Lewis is then able to define causal dependence of distinct actual events as determined by counterfactual dependence as noted by Hume's second statement: "[...] *if the first object had not been, the second never had existed*"<sup>6</sup> for it can be explained as follows:

---

<sup>2</sup> "[...] one world is close to actuality than another if the first resembles our actual world more than the second does." *Ibid*, p.163

<sup>3</sup> It is important to remark that A and C are propositions not referring to any linguistic entity, but to sets of possible worlds.

<sup>4</sup> So that if A depends counterfactually on B then if it were  $a_1, a_2, \dots, a_n$ , (where no two of which are compossible) then it would have been  $b_1, b_2, \dots, b_n$ .

<sup>5</sup> *Ibid* p.166

<sup>6</sup> Hume, *Enquiry*, Section VII

“where  $c$  and  $e$  are two distinct causal events,  $e$  causally depends on  $c$  iff  $c$  does not occur  $\square \rightarrow e$  does not occur” (being  $c$  occurs  $\square \rightarrow e$  occurs automatically true).

However a third step is still required in order to define causation from causal dependence.

Lewis considers that for every distinct event that is related by causal dependence, there is between them a relation of causation, even though causal dependence is *sufficient* but not *necessary* for causation itself<sup>7</sup>. However we can render causal dependence *transitive* by making one or more steps of causal dependence that define the *causal chain* among distinct events  $c$  and  $e$  where  $e$  does not *directly* depend on  $c$ . So that  $c$  is a cause of  $e$  iff there exists a causal chain leading from  $c$  to  $e$ .

### 1.3 \_ Conclusion

With this theory we are now able to define that an effect counterfactually depends upon its cause iff:

- All the counterfactuals  $A_1 \square \rightarrow C_1, A_2 \square \rightarrow C_2, \dots$  between corresponding propositions in the two families are true;
- The family  $e_1, e_2, \dots$  causally depends on the family  $c_1, c_2, \dots$  (So that whether  $e$  occurs or not depends on whether  $c$  occurs or not);
- There exists a causal chain leading from  $c$  to  $e$  so that its causation is explainable by transitive causal dependence.

## 2 \_ Actual Objections

### 2.1 \_ Early problems

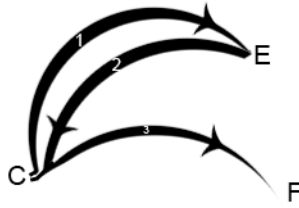
There are three major problems concerning counterfactual analysis of causal dependence.

#### 2.1.1 \_ Epiphenomena

The problem of Epiphenomena lies on the assumption that if the effect  $e$  had not occurred then its cause  $c$  would not have occurred. And, as Lewis pointed out (1973), in this way we have a *spurious reverse causal dependence*. Indeed, by considering the following example:

---

<sup>7</sup> “Causal dependence among actual events implies causation. [...] Causation must always be *transitive*, Causal dependence may not be”. Lewis (1973).



we notice that  $c$  causes first  $e$  and then  $f$ , but  $e$  does not cause  $f$ . Therefore, a proper solution (avoiding back-tracking analysis) to the possible assumption that, ‘if the epiphenomena  $e$  had not occurred then its cause  $c$  would not have occurred and the further effect  $f$  would not have occurred either’, is by denying the counterfactual itself; so that ‘if  $e$  had been absent, it is not that  $c$  would have been absent’. Rather  $c$  would have occurred failing to cause  $e$ .

For Lewis this is possible by considering the case of violations of law, for it is the only way in order to determine divergence in determinism<sup>8</sup>.

### 2.1.2 \_ Pre-emption

The problem of pre-emption establishes that two or more causes are involved in the occurrence of an event  $e$ . Two possibilities follow:

**Early pre-emption:** where the occurrence of  $c_1$  avoids at  $t_0$  the occurrence of  $c_2$ . In this case the determination of an ‘intermediate’ moment  $d$  (at  $t_1$ ) can define the causal chain that determine  $c_1$  as cause of  $e$ .

**Late pre-emption:** where the failure of  $c_2$  depends on an earlier occurrence of  $e$  caused by  $c_1$ . Here a solution cannot be found in (1973) theory, but in Lewis (1999) where is introduced the concept of *alteration* of an event such that its *fragility* depends on the time the event occurs; therefore  $e$  would have not been the same if caused by  $c_2$  (see 2.3 in this paper for proper explanation).

### 2.1.3 \_ Trumping pre-emption

In this case both  $c_1$  and  $c_2$  simultaneously determine the occurrence of an event  $e$  not letting any possibility of defining any causal chain that could show all possible intermediate events, so to determine either the first or the second cause. Also, no *alteration* of the event  $e$  occurs<sup>9</sup>. However, if we can consider that, by altering  $c_1$  while holding fixed  $c_2$  the event  $e$  would have been at least slightly different, we cannot say the same in case of  $c_2$  being altered

<sup>8</sup> Lewis, *Causation*, p.170, 171.

<sup>9</sup> For a proper understanding refer to J. Schaffer (2000).

and holding fixed  $c_1$ ; but in order to respond to some skeptics answers, Lewis is forced to determine the necessity of some *metric of distance* in alterations.

Among all the objections presented in the past 30 years to the theory of counterfactuals, I will consider the following in order to proceed in explaining Lewis's latest analysis:

2.2: Fine and the counterintuitive falsity of counterfactuals

2.3: Menzies and probabilistic theories of causation

## 2.2 \_ Fine and the counterintuitive falsity of counterfactuals

By formulating the example constituted by the relation of the events: "Nixon pressing the button" and "a nuclear war takes place", K. Fine (1975) tries to demonstrate that Lewis' concept of similarity relation is wrong for: if there will never be a nuclear holocaust, then Lewis analysis is false. In fact " given any world in which antecedent and consequent are both true, it will be easy [by following Lewis] to imagine a closer world in which the antecedent is true and the consequent is false<sup>10</sup>". This objection brings Lewis to the following clarifications:

- (i) The difference between *convergence* and *divergence* between similar worlds.
- (ii) The establishment of a hierarchy of rules for determining similarity.

(i) By using Fine's example, Lewis<sup>11</sup> formulates the constituents of four possible worlds that summarise the main possibility to determine any fundamental relation of similarity to the actual world. So that, being  $w$  our actual world, we may have a relation between  $w_1$  and  $w$  where shortly before  $t$ , due to a little (or tiny) miracle<sup>12</sup>, a divergence starts; or a relation between  $w_2$  and  $w$ , where there are no miracles happening, and Nixon presses the button. However, in this case the *convergence* miracle required to wipe out Nixon's action will always be bigger than the one required to *diverge* from  $w$  and allows Nixon to press the button. This is possible, however, only by formulating a time's arrow moving from past to future where we always have an asymmetry of counterfactual dependence between an open future and a fixed past through the actual present; and, again, an abolition of back-tracking analysis.

(ii) In analyzing the four possible worlds, Lewis can establish the following hierarchy in order to determine the closeness of possible worlds:

---

<sup>10</sup> K. Fine, *Counterfactuals*, 1975, p.452

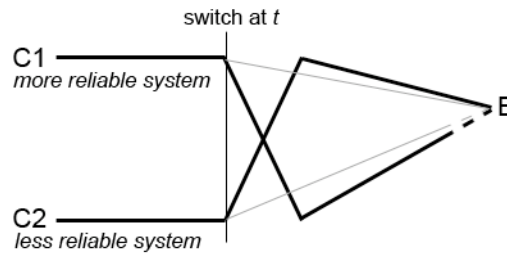
<sup>11</sup> Lewis, *Counterfactual Dependence and Time's Arrow*, p.44-48

<sup>12</sup> Miracle is used by Lewis to express a relation between different worlds where we can determine a (big or small) violation of law.

- We first try to avoid *big* violations of laws
- Secondly, we maximise the spatio-temporal region where particular facts prevail [Lewis defines in (1986b) events as classes of possible spatio-temporal regions<sup>13</sup>]
- We can then try to avoid also small violations of law
- We can finally try to secure similarity of particular facts (which is of less or no importance).

### 2.3 \_ Menzies and probabilistic theories of causation

In Menzies (1989), we face the problem of *conflict* between a pre-empting actual chance (C1) and a pre-empting potential one (C2) when applying probabilistic theory of causation<sup>14</sup> to counterfactual dependence. In short Menzies' example can be shown as follows:



whereas my “throwing the switch” shuts down the reliable system and turns on the unreliable one. Here the pre-empting actual cause decreases the chance of the effect while the pre-empted potential cause increases its chance.

In order to answer this, Lewis (1999) improves the previous theory (1973) of ‘whether’ one event occurs, with the following *whether*,

<sup>13</sup> This is due to Kim’s critics of counterfactual dependence as broader than causal dependence (1979). Briefly:

“Causal dependency is only one among the heterogeneous group of dependency relationship that can be expressed by counterfactuals”. Four problems follow:

- It is a problem of analytical dependency in defining a causal relation in counterfactual propositions where we could never admit any causal dependence at all between events.
- It is a problem in which one event is a constituent part of another. And here again, no causal relation between events is possible even though it is stated by the counterfactual.
- It is a problem of determining dependence between events when an agent does an action by doing another. Yet, the counterfactual states something that has no causal dependence at all.
- It is a problem when we have an event that asymmetrically depends on the first, but is not a causal consequence of it.

<sup>14</sup> It is not possible to analyse the problem here due to space limit.

*when* and *how* an event occurs on *whether*, *when* and *how* another event occurs such that he can define the possibility of *fragile* events that, by occurrence *at* a slightly different time and/or *in* a slightly different manner, it would have been a *different* event.

So that, if it is still considered within the boundaries of deterministic causation, we do not lose the chain that leads from *c* to *e*, no matter if they are involved into a early or a late pre-emption with another possible cause *c1*.

### **3 \_ Conclusions**

So far we have analysed the problems relating causation (1.1); its possible resolution by the contemporary analysis of counterfactuals related to causal dependence (1.2); and some derivable problems concerning the *closeness* of possible worlds and the actualization of pre-emption (2).

This last problem, however, represents a limit for developments of new theories of causation.

I therefore will concentrate this last chapter on analyzing some recent conclusions about the concept of pre-emption related to the theory of counterfactuals.

One problem resulted in defining pre-emption as an *intrinsic relation* between events, so that the potential pre-empting cause determines a sort of *quasi-dependence* between events. Menzies (1996) proposes a different approach by revealing the possibility of *commonsense intuitions* such that “*c* causes *e* iff the intrinsic relation that *typically* accompanies causal dependence holds between *c* and *e*”. However, I consider that this problem was already pointed out by Lewis (1986*b*) in relation with *sensitivity* (which I consider to be of immense importance), where by *commonsense* we define the sufficient boundaries of related causal events, such that we cannot degenerate in regressions like ‘the Big-Bang is the cause of death of every human being ever existed’<sup>15</sup>.

Indeed Lewis argues on every possibility of *quasi-dependence* in terms of counterfactual analysis, for it is not possible to reveal any causal chain between events; and as we have seen, this is a fundamental element in determining the *transitivity* of causal dependence; therefore no intrinsic determination can be applied to the solution of the problem of pre-emption.

---

<sup>15</sup> If a chain is insensitive enough that you can predict it, then it is insensitive enough that you can kill by it”. Lewis, *Postscripts to Causation*, p.187.