Lecture 4 – Lewis, Part II: Counterfactual Theory and Objections

1. Counterfactual and Causal Dependence

On Lewis's counterfactual theory, some event *Y* is causally dependent on a distinct event *X* if and only if *Y* (or, more accurately, the proposition that *Y* obtains) is counterfactually dependent on (the proposition that) *X* (obtains). And, in order for *Y* to be counterfactually dependent on *X*, the following two counterfactuals must be true:

(1) $X \mathrel{\square\!\!\rightarrow} Y$

(2) $\neg X \mathrel{\square\!\!\rightarrow} \neg Y$

Where a counterfactual is true if and only if all the closest possible *X*-worlds are *Y*-worlds.

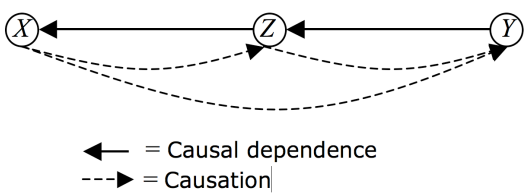
2. From Causal Dependence to Causation

Recall that causal dependence **suffices** for causation, but is **not necessary** for it. To see why, consider the following scenario.

> *Two expert marksmen are sent to assassinate a tyrant. Assassin A has been ordered not to fire unless Assassin B does not fire. Both assassins take aim, Assassin B fires and hits the target. The tyrant is killed.*

It is clearly the case that Assassin B's firing caused the tyrant's death. However, that death was *not* causally dependent on B's firing. After all, if B had not fired, A would have fired and killed the tyrant. This is a case of **preemption**.

In order to avoid counterexamples like these, Lewis defines **causation** as follows. *X* is a cause of *Y* if and only if there is a **causal chain** leading from *X* to *Y*. And such a causal chain exists if and only if there is a set of events (of at least two members) beginning with *X* and leading to *Y*, where each event is causally dependent on the previous event in the chain. When the chain only contains two events, *X* and *Y*, *Y* is caused by *X* in virtue of being causally dependent on *X*. When there are three or more events they will stand in relations as modelled below:



From this it follows that **causation is transitive**, but causal dependence is not.

NB: A relation is *transitive* if and only if, if the relation holds between *A* and *B*, and between *B* and *C*, then it also holds between *A* and *C*.

3. Counterexamples and Responses

Background Conditions

There are several different species of counterexample to Lewis's counterfactual theory. One of these concerns background conditions. Recall from the lecture on Mackie that, in the case of the house fire, oxygen turned out to be an INUS condition for the fire. Similarly, **the fire is causally dependent on the presence of oxygen**.

In response, Lewis accepts this consequence of his theory. He explains away our intuition (that the oxygen is not a *cause* of the fire) by arguing that it is informed by the **pragmatics of explanation**. The intuition arises because *saying* oxygen was a cause of the fire is often deemed inappropriate.

Transitivity

A second class of counterexamples concerns the transitivity of causation. In many cases this is quite an intuitive thing to think. But consider the following case:

> *A hiker is hiking on a path beneath a cliff when a boulder falls from the cliff overhead, heading straight towards her. Spotting the boulder, the hiker ducks, the boulder misses her, and she survives.*

The hiker's survival causally depended on the ducking, and the ducking causally depended on the boulder's falling. So **the boulder's falling is a cause of the hiker's survival**.

In response to these cases, Lewis employs the same defence as he did against counterexamples from background conditions. The strange-sounding causal claims are ***technically speaking* true, but inappropriate to say**.

Causal dependence of causes on their effects

Temperature readings on thermometers are counterfactually dependent on the air temperature in room. But, intuitively, the opposite is also true. Suppose the temperature in a room is 20ºC, and a thermometer in that room reads '20ºC'. Both of the following counterfactuals seems true (and so, on Lewis's theory the temperature turns out to be causally dependent on the thermometer reading):

(1a) $\boldsymbol{Therm_{20ºC} \mathrel{\square\!\!\rightarrow} Temp_{20ºC}}$

(2a) $\boldsymbol{\neg Therm_{20ºC} \mathrel{\square\!\!\rightarrow} \neg Temp_{20ºC}}$

(1a) is immediately true, since both antecedent and consequent actually obtain. Lewis's response to this kind of counterexample is to deny the truth of (b). He argues that $\neg Therm_{20ºC}$-worlds where $\neg Temp_{20ºC}$ obtains are **less similar** to (i.e **more distant** from) the actual world. So, (2a) is false, and counterfactual dependence fails between these two events.