

De facto dependence. Preemption seems to show that there can be counterfactual dependence without causation. But if one looks at examples of preemption, it often seems that e does in a way depend on c . It depends on c “in the obtaining circumstances” or “holding certain things fixed.” One implementation of the holding fixed strategy is the method of causal networks. Hitchcock’s version says c causes e iff there’s an “active” path from C to E through the relevant network, in the sense that changing C ’s value *with off-path variables frozen at their actual values* changes E ’s value.

This view runs into the problem of *self-countering threats*. Suppose e was going to happen “anyway,” until c came along and presented threat d , at the same time protecting e from d via (bodyguard) event b . The path $C \rightarrow B \rightarrow E$ is active, since if we hold threat D fixed at its actual value of 1, toggling C by toggling B toggles E . Eg. let C = my putting the bomb under your chair, E = your survival, D = your chair’s exploding, and B = your fleeing the room pre-explosion on noticing a bomb under your chair.

Sufficient reason (1). Distinguish a variable V ’s taking a *default* value from its taking a *deviant* value. Default values are what you’d expect anyway; deviant values depart from the norm in a way that intuitively requires explanation. The *principle of sufficient reason* says that V takes a deviant value only if one of its parents does. A network *self-contained* if every variable in the network satisfies the principle of sufficient reason. “The main idea I will defend [called TC for token causation] is that counterfactual dependence [dependence not holding anything fixed] is necessary and sufficient for token causation in self-contained networks.” The bomb network is self-contained, and there’s no counterfactual dependence. So we now have a story about why the faux hero is not a cause.

Objection. Let’s consider a homeostatic framework which naturally seeks a certain rest state -- say, an Aristotelian framework in which objects’ being at rest in their natural home is default. There’s a marble “here” in its natural home at time T . Intuitively its being here at T is a cause of its being here five minutes later. If the marble had been elsewhere, though, say, 3’ up in the air, Mother Earth would have pulled it 3’ down, and so on. Since the marble would still have wound up “here” at $T+5$ wherever it had been at T , there’s no counterfactual dependence. The model seems self-contained, since Mother Earth doesn’t act unless the marble starts out in a deviant location, away from its natural home. TC then seems to predict that predict the marble’s location at T is not a cause of its location at $T+5$, even though the marble is undisturbed over that period.

Objection. The network account’s pride and joy is standard early preemption cases, e.g., Suzy throwing so that Billy doesn’t bother. But the network here is not self-contained, for B takes a deviant value (1 for Billy-throw) when its lone parent S assumes its default

value (0 for Suzy-no-throw). This means some of our strongest positive judgments are no longer confirmed by the theory.

Sufficient reason (2). “What [if] the causal network is not self-contained? ...What I suspect is that [the active-route account] is essentially correct, but that [its] verdicts can be partially or completely counteracted by TC when the latter yields a clear negative conclusion.” So: c causes e iff EITHER the network is self-contained and e depends on c absolutely, that is holding nothing fixed (note sufficient reason is automatically satisfied), or it is not self-contained and e depends on c holding off-path variables fixed at their default values (so as to minimize violations of sufficient reason).

Objection. You can rewrite preemption cases to make the network self-contained. Maybe Suzy is bound to throw; it isn’t her failure to throw that gets Billy to throw but her throwing badly. Unless bad throws are default (?), Billy’s throwing does not violate PSR, so the network is self-contained, so TC^+ predicts no causation. Also you may be able to rewrite self-countering threat cases to make the network not self-contained, so that TC^+ wrongly predicts causation.

New theory. c is a cause of e iff e depends on c holding fixed some “good” G , that is, for some good G , $\sim Oc \ \& \ G \ \square \rightarrow \sim Oe$. Part of goodness is truth. But that’s not enough to deal with self-countering threats, as it’s true that the chair explodes. Here’s the basic idea: if c makes itself indispensable only by presenting a threat G , then c should lose all relevance in G ’s absence. E.g., if for some reason the chair doesn’t explode, then my putting the bomb under your chair does nothing to make you healthier. Say that e depends on c modulo G if (i) and c appends e modulo $\sim G$ if (ii).

- (i) in nearby worlds with c and e – these will be G -worlds – removing c holding G fixed also removes e ($\sim Oc \ \& \ G \ \square \rightarrow \sim Oe$)
- (ii) in nearby worlds omitting c and e – these will be $\sim G$ worlds – adding c while holding $\sim G$ fixed brings along e . ($Oc \ \& \ \sim G \ \square \rightarrow Oe$)

G is *good* iff it is true and we have dependence (i) and appendence (ii). Self-countering threats satisfy dependence but not appendence. In nearby worlds where the bomb is *not* put in place yet you do *not* survive, adding the bomb while holding fixed the non-explosion is not a good way of getting you to survive. Preempted causes like Billy’s throw satisfy appendence; in nearby worlds where Billy doesn’t throw and the bottle doesn’t break, adding the throw while holding fixed non-deflection is a good way of getting the bottle to break. But not dependence. Holding fixed that Billy’s rock *is* deflected, subtracting his throw does not stop the bottle from breaking. So you need both.