Part 2, Paper 1 Metaphysics | Lent 2017

Causation

Lecture 2: Counterfactuals

1. Lewis' Counterfactual Analysis

"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. Had it been absent, its effects—some of them, at least, and usually all — would have been absent as well." (Lewis, 1973, p.161)

In it's simple form, a counterfactual theory of causation says that:

> c causes e iff: c and e are distinct events, and if c were not to occur, then e would not occur either.

Lewis' general semantics of counterfactuals: 'if c were not to occur, then e would not occur' ($\sim c \mathrel{\square\!\!\rightarrow} \sim e$) is true if some possible world without c in which e doesn't occur either is more similar to the actual world than any possible world without c in which e does occur.

2. Symmetry and Backtracking

Lewis (1979) maintains that the asymmetries of causation can be explained in terms of the asymmetry of counterfactual dependence.

Can Lewis's theory explain that effects counterfactually depend on their causes but not *vice versa*? This would require counterfactual dependence to be an asymmetric relation. It is not obviously so. Consider (C) my ordering a book on Amazon yesterday, and (E) its being delivered at noon today.

1. If I hadn't ordered the book yesterday, it would not have been delivered at noon.
2. If it hadn't been delivered at noon, then it would have to be that I didn't order it yesterday.

We can make sense of both as being true counterfactual conditionals.

Lewis' solution is to disqualify such *backtracking* counterfactuals: (1) is true, while (2) is false. This is because the most similar worlds in which the book was not delivered are those in which:

- History is the same until just before noon
- A small miracle occurs just before noon (perhaps, unlike what actually happened, the delivery person lost my book at 11am)
- Everything obeys the laws of nature thereafter

3. Defending this picture

If Lewis is right, then the relevant kind of counterfactual dependence is asymmetric, and moreover time-assymetric. But why should we accept Lewis' stipulations of which worlds are most similar to the actual world?

One option: make the similarity relation one according to which by definition, for any possible world, worlds preserving the past are always more similar overall than worlds not preserving the past.

Lewis's answer is an empirical one. The asymmetry of counterfactual dependence (in the actual world) arises because of an asymmetry of miracles, which in turn is due to the asymmetry of overdetermination: earlier events are massively overdetermined by later events, but not vice versa.

If overdetermination is indeed asymmetric in this way, then it is typically true that events causally depend on earlier events but not on later events. Lewis allows this to be no more than a contingent fact. Backwards or time-reversed causation may stil be a possibility that should not be ruled out *a priori*. Notice how here Hume's 'semantic convention' now becomes an empirical generalisation.

4. Problems with this explanation

Psychological Implausibility: As account of the cause-effect relation it is psychologically implausible (Horwich 1987)

Asymmetry of Miracles is False: Lewis' asymmetry of miracles does not hold in many cases (Elga 2000)

Physical Incorrectness: The asymmetry of overdetermination does not exist in the form required to support Lewis's explanation (Price and Westlake 2009)

References

Elga, Adam, 2000. "Statistical Mechanics and the Asymmetry of Counterfactual Dependence", *Philosophy of Science*, 68 (Supplement): 313–324.

Horwich, Paul, 1987. *Asymmetries in Time*, Cambridge, Mass: MIT Press.

Lewis, David. 1973. "Causation", *Journal of Philosophy*, 70: 556–67.

Price, Huw, and Brad Weslake, 2009. "The Time-Asymmetry of Causation", in Beebee, Hitchcock,and Menzies, *The Oxford Handbook of Causation*, Oxford: Oxford University Press, pp. 414–43.

Maarten Steenhagen (ms2416@cam.ac.uk) | February 2017