Case Study 3: The Analysis of Causation

1. Lewis’ Methodology
In the introduction to his *Philosophical Papers, Vol. I*, Lewis provides a helpful summary of his general philosophical methodology (p. x):

> Our “intuitions” are simply opinions; our philosophical theories are the same. Some are commonsensical, some are sophisticated; some are particular, some general; some are more firmly held, some less. But they are all opinions, and a reasonable goal for a philosopher is to bring them into equilibrium. Our common task is to find out what equilibria there are that can withstand examination, but it remains for each of us to come to rest at one or another of them. If we lose our moorings in everyday common sense, our fault is not that we ignore part of our evidence. Rather, the trouble is that we settle for a very inadequate equilibrium.

Thus most of the time Lewis talks about the verdicts of common sense, rather than the verdicts of intuition.

2. Regularity Analyses of Causation
Before the advent of counterfactual analyses in the 1970’s, so-called regularity analyses were the predominant approach to analyzing causation.

The basic idea behind a regularity analysis: event $c$ causes event $e$ iff the fact that $c$ occurs, together with the fundamental laws of nature and some suitably chosen summary of the relevant background conditions, entails that event $e$ occurs.

Some problems for such an account:

- **the problem of epiphenomena**: If event $c$ causes both event $e_1$ and event $e_2$, it might be that $e_1$ could not, given the laws and the relevant background conditions, have occurred without $e_2$ also occurring.

- **the problem of preemption**: Suppose that $c_1$ causes $e$, but if it hadn’t, then $c_2$ would have caused $e$ instead: $c_2$ is a preempted potential cause of $e$. In most cases of this sort, it will be the case that $c_2$ could not, given the laws and the relevant background conditions, have occurred without $e$ also occurring.

3. A Crash Course on Counterfactuals
So-called counterfactual (or subjunctive) conditionals are conditional sentences in the subjunctive mood:

- indicative conditional: “If A is true, then C is true” (in symbols: “$A \supset C$”)
- counterfactual conditional: “If A were true, then C would be true” (in symbols: “$A \Box \rightarrow C$”)

Let an A-world be a possible world in which A is true. Then here are two accounts of the truth-conditions for counterfactual conditionals (ignoring cases in which A is necessarily false):

*Stalnaker’s proposal*: “$A \Box \rightarrow C$” is true iff, in the A-world most similar to the actual world, C is true.

*Lewis’ proposal*: “$A \Box \rightarrow C$” is true iff some A-world where C holds is more similar to the actual world than any A-world where C does not hold.

Why might we prefer Lewis’ proposal over Stalnaker’s? Stalnaker’s proposal breaks down if (i) there are several most similar A-worlds, or (ii) there is an infinite, unbounded sequence of more and more similar A-worlds.

*Note 1*: Both accounts analyze counterfactuals in terms of possible worlds plus degree of overall similarity between these worlds and actuality.

*Note 2*: On both accounts, if A is true, then “$A \Box \rightarrow C$” entails “$A \supset C$.”

For our purposes, it will often be easier to evaluate counterfactuals in terms of Stalnaker’s original proposal.
4. Lewis’ Original Analysis of Causation

Let us define a notion of **counterfactual dependence** among events as follows:

Event \( e \) counterfactually depends on distinct event \( c \) iff:

1. \( \square \neg O(c) \rightarrow \\
   \neg O(e) \) is true, and

2. \( \neg O(c) \square \\
   \neg O(e) \) is true.

However, since \( O(c) \) and \( O(e) \) together entail \( \square \neg O(c) \rightarrow O(e) \), we can rewrite this definition as follows:

If distinct events \( c \) and \( e \) both occur, then \( e \) counterfactually depends on \( c \) iff \( \neg O(c) \square \\
\neg O(e) \) is true.

Lewis then defines causation to be the ancestral of counterfactual dependence:

\[ \text{Lewis’ original analysis of causation: Event } e \text{ is a cause of event } c \text{ iff:} \]

1. \( c \) and \( e \) both occur,
2. \( c \) and \( e \) are distinct, and
3. there is a (possibly empty) set of distinct events \{\( d_1, d_2, \ldots, d_n \)\} such that \( d_1 \) counterfactually depends on \( c \), \( d_2 \) counterfactually depends on \( d_1 \), \ldots, and \( e \) counterfactually depends on \( d_n \).

How this analysis handles some of the problems for regularity analyses of causation:

- **the problem of epiphenomena**: If we rule out so-called backtracking readings of the relevant counterfactual conditionals (readings where we talk about how things would have been different in the past, if they had been different now), the problem disappears.

- **the problem of preemption**: For some kinds of preemption (so-called early cutting preemption), the ban on backtracking readings of the relevant counterfactuals is enough to dispel the problem.

5. Problems for Lewis’ Original Analysis

Some difficulties one faces when filling in the details of Lewis’ original analysis:

- How do we determine the relevant similarity relation between worlds?
- What are we to consider when we consider what would happen if a given event did not occur?
- How exactly do we individuate events?

Potential counterexamples to Lewis’ original analysis:

- **symmetric overdetermination**: Events \( c_1 \) and \( c_2 \) both occur and jointly bring about event \( e \), yet either \( c_1 \) or \( c_2 \) on its own would have been sufficient to bring about \( e \).

- **late cutting preemption**: “Billy and Suzy throw rocks at a bottle. Suzy throws first, or maybe she throws harder. Her rock arrives first. The bottle shatters. When Billy’s rock gets to where the bottle used to be, there is nothing there but flying shards of glass. Without Suzy’s throw, the impact of Billy’s rock on the intact bottle would have been one of the final steps in the causal chain from Billy’s throw to the shattering of the bottle. But, thanks to Suzy’s preempted throw, that impact never happens” (“Causation by Influence,” p. 184).

- **trumping preemption**: “The sergeant and the major are shouting orders at the soldiers. The soldiers know that in case of conflict, they must obey the superior officer. But as it happens, there is no conflict. Sergeant and major simultaneously shout ‘Advance!’; the soldiers hear them both; the soldiers advance. Their advancing is redundantly caused: if the sergeant had shouted ‘Advance!’ and the major had been silent, or the major had shouted ‘Advance!’ and the sergeant had been silent, the soldiers would have still advanced. But the redundancy is asymmetrical: since the soldiers obey the superior officer, they advance because the major orders them to, not because the sergeant does” (“Causation by Influence,” p. 183).
6. Lewis’ Emended Analysis of Causation

Some vocabulary:

Event e is fragile if, or to the extent that, it could not have occurred at a different time, or in a different manner.

An alteration of event e is either a very fragile version of e, or else a very fragile alternative event that is similar to, but numerically distinct from, e.

Then instead of his old notion of counterfactual dependence, Lewis instead defines a new notion of influence:

If c and e are distinct actual events, then c influences e iff there is a substantial range c₁, c₂, . . . of different not-too-distant alterations of c (including the actual alteration of c) and there is a substantial range e₁, e₂, . . . of not-too-distant alterations of e (at least some of which differ), such that “O(c₁) □→ O(e₁)” is true, “O(c₂) □→ O(e₂)” is true, and so on.

As before, Lewis defines causation to be the ancestral of influence:

Lewis’ emended analysis of causation: Event c is a cause of event e iff:
1. c and e both occur,
2. c and e are distinct, and
3. there is a (possibly empty) set of distinct, actually occurring events {d₁, d₂, . . . , dₙ} such that c influences d₁, d₁ influences d₂, d₂ influences d₃, . . . , and dₙ influences e.

7. Problems for the Emended Analysis

Two potential problems for Lewis’ emended analysis:

- transitivity: By taking causation to be the ancestral of influence, Lewis has ensured that, on his account, causation is always transitive. However, there seem to be certain intuitive counterexamples to the transitivity of causation, including the following:

  Black and Red: “Imagine a conflict between Black and Red. . . . Black makes a move that, if not countered, would have advanced his cause. Red responds with an effective countermove, which gives Red the victory. Black’s move causes Red’s countermove, Red’s countermove causes Red’s victory. But does Black’s move cause Red’s victory? Sometimes it seems not” (“Causation by Influence,” p. 194).

  Lewis’ response is to try to explain away the intuition that Black’s move doesn’t cause Red’s victory. Two reasons why intuition might be confused about this case:

    1. We might be confusing what is generally conducive to what with what caused what. (After all, moves like Black’s are generally conducive to Red losing.)
    2. We might be confusing whether-whether dependence with causation. (After all, Red would have won whether or not Black made his move.)

- causation by omission: We often talk of omissions being both causes and effects. However, omissions are not events, so how do we extend Lewis’ account to allow for causation by omission? Lewis’ solution is first to extend his term “alteration” so that it pertains to both events and omissions, and then to replace “event” with “event or omission” everywhere in his analysis of causation.

  This proposal allows Lewis to countenance causation by omission, but it also yields the result that there is much, much more causation by omission than we intuitively deem there to be.

  Lewis’ response: even though there really is such widespread causation by omission, for pragmatic reasons we don’t usually mention it in conversation: “There are ever so many reasons why it might be inappropriate to say something true. It might be irrelevant to the conversation, it might convey a false hint, it might be known already to all concerned . . .” (“Causation by Influence,” p. 196).
8. Appendix: Ways to Resist Recalcitrant Intuitions


1. Insist that our intuitions are not really so firm, so the case is an example of “spoils to the victor.”
2. Insist that the case is misleading, because it is easily confused with other, unproblematic cases.
3. Insist that the case is misleading, because it is misleadingly presented.
4. Insist that the example is misleading, because it naturally draws our attention to other, related notions which we easily confuse with the notion under dispute.
5. Argue that the recalcitrant intuitions conflict with sacrosanct general principles about the notion in question, and so ought to be dismissed.
6. Argue that the unintuitive verdict issued by the theory isn’t false—it’s just odd to say, for pragmatic reasons.
7. Insist that intuitions about the case are too easily buffeted about to be taken seriously.
8. Insist that the case is too outré for intuitions about it to be of much concern.