Philosophy 125 — Day 27: Overview

- Branden's Office Hours This Week: W 2–4 & F 2:30–4:30 (not T). Vanessa's: Th 11:15–1:15. Josh's: F 3–5. Branden's Next Week: T 2–4 & W 4–6.
- Final Exam Questions Have Been Announced (see website / email)
- See Last Lecture Notes (Slide #2) for Hints on Study Questions 2 and 3
- Plan: One last lecture (mercifully) today, then review + evaluations Thursday
- Two New Links: Vranas on Time-Travel, Hall & Paul on Counterexamples
- A Great Talk this Afternoon in Statistics (or why I rescheduled today's OH's)
  - "A statistician flips a coin" – Persi Diaconis (Statistician, Mathematician, and Magician Extraordinaire) – 4 p.m. Today at 1011 Evans Hall
- Agenda: Causation, Finalé
  - What distinguishes causal from non-causal sequences (connection)? (last time)
  - What distinguishes causes from effects in causal sequences (direction)? (today)
  - What distinguishes causes from conditions in causal seqs (selection)? (today)
  - Epilogue on the paradoxes of time travel (time permitting)

Introduction to Causation III: What *is* the causal *relation*? 4

- Most people think causation is *asymmetric*. But, what distinguishes causes from effects? That is, what determines the *direction* of the causal relation?
- There are six main arguments concerning the direction of causation. Three of these arguments say that causal direction matches temporal direction. Three of these arguments say that causal direction does not match temporal direction.
- **Bilking**: The main argument that causal order matches temporal order is the bilking argument (Black). A drawing made by an alleged clairovoyant on Monday might be caused be a pattern made Tuesday. But, as Mackie explains:
  > On every occasion, after the drawing is made, it is possible that someone or something should intervene so that the corresponding pattern fails to be produced . . . [So] it cannot on any occasion be the pattern that is responsible for the details of the drawing: the precognition hypothesis must be false even for those occasions when the device is not stopped, when the pattern is actually produced and turns out to be just like the drawing.

  ∴ Backwards causation is impossible. There are two main replies to bilking.

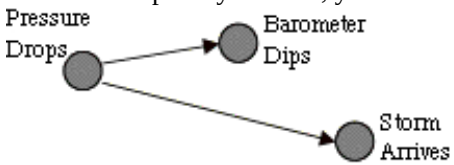
- *Intervention Restriction* (Dummett): The bilking argument only applies to cases in which human intervention is possible. What is there to prevent backwards causation when human intervention is ruled out?
- *Determinism/Indeterminism Dilemma*: If the world is deterministic, then the bilking intervention is impossible, as it will already be fixed that the later cause will occur. If the world is indeterministic, then the bilking intervention is possible but no longer problematic, as the case will then be one in which the earlier event (*e.g.*, the clairvoyant's drawing) is an uncaused indeterministic eruption. [Is this second horn right? What about indeterministic causation?]

- **Time Travel**: The first main argument *against* the causal order matching the temporal order is that backwards causation *is* possible in cases of time travel.
- It seems metaphysically possible to enter a time machine at t_1, thus causing oneself to exit the time machine at an earlier time t_0. Gödel found solutions to Einstein's field equations that permit such loops: "By making a round trip on a rocket ship it is possible to travel into any region of the past, present, and future and back again, exactly as it is possible in other worlds to travel into distant parts of space."

- There are three main replies to the time travel argument.
- *Time travel is incoherent.* A variety of incoherencies are alleged, including (*i*) the incoherency of changing what is already fixed (the past), of (*ii*) being both able and unable to kill one's own ancestors, or of (*iii*) generating a causal loop and a reflexive relation of "self-causation". Gödel solutions may be dismissed as mathematical artifacts, not reflecting any possible situation [it's unclear whether time-travel is incoherent: see Vranas' paper for penetrating analysis].
- *Local Direction.* Time travel involves *locally forward* causal steps. In Gödel solutions, spacetime is structured in such a way that a series of locally forward steps produces a globally backwards path. This is compatible with the causal order being the temporal order, at least (locally) at each particular step. In General Relativity, it's not always meaningful to talk about "*the* past".
- *Redescription.* Instead of the time traveler entering the machine at t_1 and exiting at t_0, this may be redescribed in terms of the spontaneous creation at t_0 of one individual, and the spontaneous disappearance at t_1 of another, with coincidental correlations between them (different people, so no paradoxes).

- **Simultaneous Causation**. Second main argument against the causal order being the temporal order: simultaneous causation is possible, *e.g.*, when an iron ball depresses a cushion (Kant, Taylor, Brand). The main reply is:
- *Misdescription* (Mellor). The iron ball takes time to depress the cushion – all bodies take time to "communicate their motions". There are no perfectly rigid bodies, at least not in any nomologically possible world. Question: since when is nomological possibility the guide to the metaphysics of causation? Recall that counterexamples involving magic are not ruled-out for this reason.
- **Argument from Physics**. This is the third main argument against the causal order being the temporal order. Physicists have entertained a variety of theories entailing backward causation [*e.g.*, Feynman's theory of positrons as electrons moving backwards in time, and 'quantum handshake' explanations of quantum phenomena]. These were all serious physical hypotheses (at least, at one time) – so they "might have been true" (Horwich, Dowe). Two replies:
- *Dismissal*. Being a physicist is no barrier to incoherence. As before, a defender of temporal order may hold that forward redescription is possible.

- *Overextension*. What sort of backward causal relation do these theories entail? The 'quantum handshake' model postulates a backwards causal arrow that *seems* to be neither the tine of a past-open fork, nor a handle to manipulate the past. What, then, is the motivation for thinking of this as a *causal* relation?
- **Joint Effects**. If the fall in atmospheric pressure at t_0 causes both the dip in the barometer at t_1 and the storm at t_2, then the dip in the barometer and the storm are causally connected and temporally ordered, yet this is not causation:



- *Additional Test*. Add a test for a joint effect structure, such as the screening-off test (Reichenbach). Then, causal direction is *unscreened* temporal direction.
- *Direct Connection*. If one can [somehow] identify the *direct* connections and apply the temporal order to these *only*, one could match the diagram. *How*?
- **Causal Theory of Time**. Temporal order is determined *by* causal order *not vice versa*. Response: *denial* – time's arrow is determined by *something else*.

Introduction to Causation III: What *is* the causal *relation*? 5

- **Last Question**. What is the basis for causal selection? What distinguishes causal sequences from "mere causal factors or conditions" in a causal nexus.
- **Caprice**. *There is no objective basis for causal selection*. Mill (1846):

Nothing can better show the absence of any scientific ground for the distinction between the cause of a phenomena and its conditions, than the capricious manner in which we select from among the conditions that which we choose to denominate the cause.

- Mill's argument has been *very* influential. David Lewis (1986) concurs:

We sometimes single out one among all the causes of some event and call it 'the' cause, as if there were no others. Or we single out a few as the 'causes', calling the rest mere 'causal factors' or 'causal conditions' . . . We may select the abnormal or extraordinary causes, or those under human control, or those we deem good or bad, or just those we want to talk about. I have nothing to say about these principles of invidious discrimination.

- Thus selection is now generally dismissed as groundless, and theorists seek to isolate some "pre-selected" conception of causation. Some Replies to Caprice:

- *Predictability*. The main argument against Caprice holds that our selections are too predictable to be without an objective basis. Hart and Honore write:

In most cases where a fire has broken out the lawyer, the historian, and the plain man would refuse to say that the cause of the fire was the presence of oxygen, though no fire would have occurred without it: they would reserve the title of cause for something of the order of . . . the dropping of a lighted cigarette . . . In making this distinction . . . our choice, though responsive to the varying context of the particular occasions, is not arbitrary or haphazard.

- But what could this distinction between causes and conditions be? Ducasse maintains that it is between sufficient causes and necessary conditions:

. . . 'cause' is contrasted with 'condition' in a serviceable . . . manner: The cause of a phenomenon is a change in its antecedent circumstances which was sufficient to bring it about. A condition of a phenomenon . . . is a change, or more frequently a state, of its antecedent circumstances which was necessary to its having occurred when it did.

- But how does this capture our selection of the cigarette-dropping over the presence of oxygen? Each factor seems necessary and neither sufficient.
- Hart and Honore maintain that abnormal situations and free actions are causes, while normal situations and non-agential factors are conditions:

In distinguishing . . . causes and conditions two contrasts are of prime importance. These are the contrasts between what is abnormal and what is normal in relation to any given thing or subject matter, and between a free deliberate action and all other conditions.

- This seems to help in the case of the cigarette-dropping (abnormal) *vs* the presence of oxygen (normal), but this doesn't seem clearly distinct from the Caprice view (how, exactly, are these considerations *not* capricious?).
- *Inseparability*. A further argument against the Caprice view is that we have no concept of causation without selection. Hart and Honore explain that:

> The contrast of cause with mere conditions is an inseparable feature of all causal thinking, and constitutes . . . the meaning of causal expressions . . . .

- The upshot is that the Caprice view deprives us of any intuitive grasp on the notion of cause. How are we to judge whether or not certain cases, such as the problem cases reviewed above or any others, involve causation or not, if our judgments are infected with a component of unsystematic caprice?
- Lewis feels the pull of this problem, and retreats to a 'prior' question:

I am concerned with the prior question of what it is to be one of the causes (unselectively . . .). My analysis is meant to capture a broad and nondiscriminatory concept of causation.

- But, it is not obvious that our intuitions about causation can provide any evidence concerning this 'broad and nondiscriminatory concept' (if there be such a concept), if our intuitions are shot through with selection effects.
- **Adicity Revisited**. Perhaps additional causal relata can reconcile caprice and predictability, and explain inseparability to boot. Schaffer suggests that:

Different speakers, in different conversational contexts, will disagree about 'the cause.' If one does not know what inquiry a speaker is pursuing, one may well find her selections capricious. What is predictable about selection is that, once conversational context is fixed, one can expect widespread agreement about the cause. If one knows what inquiry a speaker is pursuing, one will find her selections predictable. This might suggest that what is varying capriciously is which contrasts are in play in a given inquiry, and what is predictable is what counts as the cause relative to the contrasts in play.

- Mackie (he is a binary theorist, but maybe he should be a ternary theorist!) speaks of the "causal field" with reference to which a causal selection is made:

A causal statement will be the answer to a causal question, and the question 'What caused this explosion?' can be expanded into 'What made the difference between those times, or those cases, within a certain range, in which no such explosion occurred, and this case in which an explosion did occur?' Both causes and effects are seen as differences within a field; anything that is part of the assumed (but commonly unstated) description of the field itself will, then, be automatically ruled out as a candidate for the role of cause.

- Schaffer's quaternary approach, as an alternative to Mackie's binary approach:

The quaternary theorist may borrow Mackie's notion of a causal field, understood now as the reservoir of implicit contrasts. So whereas Mackie takes the causal field to be a set of assumed conditions pragmatically superimposed on a pre-selective notion of connection, the quaternary theorist relocates the field to determining the semantic value of the causal alternative and the effectual difference. Causal selection . . . is a reflection of which alternatives are in question. . . . selection is determined by the values of the contrast relata. . . . there is ultimately no difference between connection and selection. The very notion of a causal connection is only well defined in light of selective contrasts.

- Don't we still need a "causal field", even on the quaternary view? That is, don't we still need to say c_1 rather than c_2 causes e_1 rather than e_2 *in field* F?

Epilogue on the Paradoxes of Time Travel

- Consider the following scenario, which seems consistent with General Relativity (I borrow this story from Vranas' paper – with a slight change):

A wormhole is . . . analogous to a tunnel: it is a shortcut connecting two otherwise distant spacetime regions. A traversable wormhole is suitable for travel by humans . . . I enter one of the two 'mouths' . . . of the wormhole in Boston, I travel through the tunnel for two minutes (as measured by my watch), and I emerge from the second wormhole mouth in Detroit. Now recall the twin paradox from special relativity: the twin who goes on a space trip and returns ages less than the twin who stays on Earth. Similarly, if one takes the Boston mouth on a space trip and returns it to Boston, then time inside the mouth will run behind Boston time. Here is how to use this effect for time travel. I enter the Boston mouth at noon, Boston time. Inside the mouth it's earlier, say 8am. I traverse the wormhole in two minutes and I emerge from the Detroit mouth at 8:02am. Then I catch a 9am flight to Boston and I arrive at 11am, in time to find my younger self asleep in bed, prior to entering the Boston mouth.

- So far, so good. Now the fun begins. What if I bring a gun with me on my trip – can I kill my sleeping younger self (YS) before he boards the Boston mouth?

- Many people have argued that I cannot kill YS – that retro-suicide is impossible. There are various arguments for this. Here is Vranas' tale:

Look at this helpless boy, lying asleep within the range of my loaded gun! Of course it's physically possible that I kill this boy; what's physically impossible is the conjunction of the propositions that I kill this boy *and* that this boy is an earlier [temporal] stage of mine – or so a proponent of the standard [story] might claim. In reply I grant that one can refer to YS by pointing to him and uttering 'this boy' (and thus without mentioning that he is an earlier stage of mine) but I deny that the physical possibility of the proposition that I kill this boy is obvious. The physical possibility of this proposition amounts to the existence of a physically possible world in which I do kill this boy. But in any such world this boy is YS (given that 'this boy' and 'YS' are rigid designators), so what happens after I kill him – does he rise from the dead and grow up to become me? If resurrection is physically impossible, then apparently so is my killing YS. . . . [but] the story is more complicated (maybe in some world YS – and thus this boy – is not an earlier stage of mine) . . .

- As Vranas explains (in great detail, and with great skill), whether retro-suicide is physically possible trades on whether resurrection physically possible, and whether YS must [nomologically] be an earlier temporal stage of mine . . .

- **Origin Essentialism** (OE). If a person-stage is a descendant of (*i.e.*, originates from) a given sperm and egg, then it is (metaphysically) impossible that this person-stage exists without being a descendant of the given sperm and egg.
- Given that I and YS are descendants of the same sperm s and egg e, OE entails that I and YS could not have coexisted without being descendants of s and e.
- Technically, this doesn't quite get us the claim that it is physically impossible that I coexist with YS without being a later stage of YS. But, we can tweak the example to make YS an *essential descendant* of mine (*e.g.*, s or e).
- Thus, if (OE) is true, and if resurrection is physically impossible, then it seems to follow that retro-suicide (suitably understood) is also impossible.
- Note: Vranas rightly points out that arguments based on the "impossiblity of changing the past" are weak, since they make strong assumptions about what it means to "change the past". If "changing the past" just means "actualizing a non-actual past", then we *can* change the past, although we *will not*. [So, hard determinism, which says "S *can* ϕ" and "S *does not* ϕ" are *inconsistent*, is also sufficient to rule-out the possibility of retro-suicide, but trivially so.]